



Durham E-Theses

Perceived Depth Control in Stereoscopic Cinematography

SUN, GENG

How to cite:

SUN, GENG (2012) *Perceived Depth Control in Stereoscopic Cinematography*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/3458/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Perceived Depth Control in Stereoscopic Cinematography

Geng Sun

A Thesis presented for the degree of
Doctor of Philosophy



Visualisation Laboratory
School of Engineering and Computing Sciences
University of Durham
England

September 2011

Dedicated to

My mum, dad, and wife

Perceived Depth Control in Stereoscopic Cinematography

Geng Sun

Submitted for the degree of Doctor of Philosophy

September 2011

Abstract

Despite the recent explosion of interest in the stereoscopic 3D (S3D) technology, the ultimate prevailing of the S3D medium is still significantly hindered by adverse effects regarding the S3D viewing discomfort. This thesis attempts to improve the S3D viewing experience by investigating perceived depth control methods in stereoscopic cinematography on desktop 3D displays. The main contributions of this work are: (1) A new method was developed to carry out human factors studies on identifying the practical limits of the 3D Comfort Zone on a given 3D display. Our results suggest that it is necessary for cinematographers to identify the specific limits of 3D Comfort Zone on the target 3D display as different 3D systems have different ranges for the 3D Comfort Zone. (2) A new dynamic depth mapping approach was proposed to improve the depth perception in stereoscopic cinematography. The results of a human-based experiment confirmed its advantages in controlling the perceived depth in viewing 3D motion pictures over the existing depth mapping methods. (3) The practicability of employing the Depth of Field (DoF) blur technique in S3D was also investigated. Our results indicate that applying the DoF blur simulation on stereoscopic content may not improve the S3D viewing experience without the real time information about what the viewer is looking at. Finally, a basic guideline for stereoscopic cinematography was introduced to summarise the new findings of this thesis alongside several well-known key factors in 3D cinematography. It is our assumption that this guideline will be of particular interest not only to 3D filmmaking but also to 3D gaming, sports broadcasting, and TV production.

Declaration

The work in this thesis is based on research carried out at the Visualisation Laboratory, School of Engineering and Computing Sciences, University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text.

Chapter 4 is based on research done in collaboration with Dr. Nick Holliman published in [Sun and Holliman, 2009].

Copyright © 2011 by Geng Sun.

“The copyright of this thesis rests with the author. No quotations from it should be published without the author’s prior written consent and information derived from it should be acknowledged”.

Acknowledgements

First of all, I would like to express my heartfelt gratitude to my parents, Jiansheng Sun and Jianying Geng for being my best role models and giving me your unending support and bottomless love. Thank you both very much, your help in getting me to where I am has been truly amazing. I would also like to thank my grandma, uncle, aunt, and my hometown Shihezi for providing me a safe haven whenever I need to seek comfort and peace.

I am very grateful to my supervisor Dr. Nick Holliman for his advice and support during this PhD.

Through my years in Durham I have met some wonderful people and made many great friends, please forgive me if I miss anyone. Djamel and Paul, thanks for being great colleagues, house-mates, and friends. I still miss living in that freezing house with you. I would also like to thank all the people I knew through pool and poker, especially Jamie, Ian, Brian, and Nathan. Winning the DUPL post-season trophy was one of the highlights of my time here. Thank you to the college staff: Sarah Lee, Caroline, Liam, and the porters. You made my living in college a very enjoyable experience. Thanks are also due to the lovely city of Durham for playing such a great host to me for all these years. You have indeed become my second home.

Finally and most importantly, I'd like to thank my wife: Yao Cheng. Meeting you 15 years ago was the best thing that ever happened to me. You've supported me through thick and thin; helped me through every difficult moment. It has been an absolute blessing for me to have your love and dedication. Thank you and love you.

Contents

Abstract	iii
Declaration	iv
Acknowledgements	v
1 Introduction	1
1.1 The Renaissance of the Stereoscopic Cinema	1
1.2 The Research Problem	2
1.2.1 The Vergence-Accommodation Conflict	2
1.2.2 3D Comfort Zone	4
1.3 Research Methods and Contributions	5
1.3.1 Limits of the 3D Comfort Zone	5
1.3.2 Depth Mapping Methods	5
1.3.3 Depth of Field Blur Technique	6
1.4 Thesis Organisation	7
2 Background	10
2.1 Human Visual System	10
2.1.1 The Human Eye	10
2.1.2 Binocular Depth Perception in the Natural World	11
2.1.3 Binocular Depth Perception in Electronic Stereoscopic Images	15
2.1.4 Other Depth Cues	19
2.1.5 Benefits of the Binocular Vision	22
2.2 3D Displays	23

Contents	vii
2.2.1 Stereoscopic Displays	23
2.2.2 Auto-stereoscopic Displays	28
2.2.3 Volumetric Displays	35
2.2.4 Holographic Displays	37
2.2.5 Head-mounted Displays	39
2.3 Summary	41
3 Stereoscopic Cinematography	42
3.1 Cinema Technologies	43
3.1.1 Review of Cinema Technology Evolution	43
3.1.2 Stereoscopic Cinema	47
3.2 Technical Issues of the Stereoscopic Cinema	53
3.2.1 Vergence-Accommodation Conflict	53
3.2.2 Camera Configuration	54
3.2.3 Ghosting and Crosstalk	57
3.2.4 Stereoscopic Window Violation	58
3.2.5 Perceptual Distortions and Multi-rigging Technique	60
3.3 Cinematic Storytelling in Stereoscopic Cinema	62
3.3.1 Cinematic Storytelling Techniques	62
3.3.2 Cinematic Storytelling in S3D	66
3.4 Summary	68
4 Previous Works on Perceived Depth Control in S3D	70
4.1 Limits of the 3D Comfort Zone	70
4.1.1 Williams and Parrish's Results	71
4.1.2 Yeh and Silverstein's Study	72
4.1.3 Woods <i>et al</i> 's Experiment	74
4.1.4 Jones <i>et al</i> 's Results	74
4.2 Existing Depth Mapping Algorithms	75
4.2.1 Ware <i>et al</i> 's Algorithm	76
4.2.2 Wartell's Algorithm	78
4.2.3 Jones <i>et al</i> 's and Holliman's Algorithms	79

4.3	Depth of Field Blur Simulation in S3D	81
4.3.1	Yano <i>et al</i> 's Experiment	82
4.3.2	Blohm et al's Study	83
4.4	Summary	85
5	Investigation and Evaluation of Perceived Depth Control Methods	87
5.1	Investigation of Different Perceived Depth Control Methods	88
5.1.1	Real-Eye Configuration	89
5.1.2	Fixed Depth Mapping Approach	90
5.1.3	New Dynamic Depth Mapping Approach	93
5.1.4	Depth of Field Blur Simulation	98
5.2	Experiment	100
5.2.1	Protocol	100
5.2.2	Results and Analysis	107
5.3	Conclusion	113
6	Identifying the 3D Comfort Zone on a Desktop 3D Display	114
6.1	Random Dot Stereogram	114
6.1.1	Random Dot Autostereogram	115
6.1.2	Generating RDS Images with the Stencil Buffer in OpenGL .	116
6.2	Experiment	120
6.2.1	Protocol	120
6.2.2	Stimulus	124
6.3	Results and Analysis	125
6.3.1	Statistical Analysis Method	126
6.3.2	Depth plane analysis based on mistakes	127
6.3.3	Depth Plane Analysis based on time	130
6.3.4	Comparison between Square and Triangle based on mistakes .	133
6.3.5	Comparison between Square and Triangle based on time . . .	135
6.4	Conclusion	138

7	Discussion	140
7.1	Limits of the 3D Comfort Zone	140
7.2	Dynamic Depth Mapping vs. Fixed Depth Mapping	142
7.3	Depth of Field Blur Simulation	144
7.4	Limitations and Applications of this Study	145
7.5	A Guideline for Stereoscopic Cinematography	147
8	Conclusions and Future Work	150
8.1	Summary	150
8.2	Future Work	153
8.2.1	Human Factors Studies	153
8.2.2	Depth Mapping in Stereoscopic Cinematography	154
8.2.3	Depth of Field Blur Simulation	154
	Bibliography	156
	Appendix	178
A	Dynamic Depth Mapping Method	178
A.1	Initiating the Frame Buffer Object	178
A.2	Updating Scene Boundaries with the Z-Buffer	179
A.3	Off-screen Rendering	180
B	Generating RDS Images with the Stencil Buffer	182
B.1	Drawing a RDS image	182
B.2	Rendering Stereoscopic RDS images with the Stencil Buffer	183

List of Figures

1.1	Numbers of 3D movies released since the year 2000 [3dm, 2011]	2
1.2	Vergence-Accommodation Conflict in S3D viewing	3
1.3	Illustration of the 3D Comfort Zone	4
1.4	Depth Mapping Method	6
1.5	The DoF Simulation	7
2.1	Optical model of the human eye [Deering, 2005]	11
2.2	Illustration of the Panum's Fusion	12
2.3	Stereo acuity defines the smallest perceptible depth difference	13
2.4	Binocular view in electronic stereoscopic images	16
2.5	Perceived depth behind (top) and in front (bottom) of the display	17
2.6	Screen Disparities	19
2.7	Monocular depth cues available in a 2D image	20
2.8	Organisation of the Section	24
2.9	Stereoscopic Displays [Holliman, 2005]	25
2.10	Anaglyph Method	25
2.11	Auto-stereoscopic Displays [Holliman, 2005]	29
2.12	Viewing freedom in an auto-stereoscopic display	30
2.13	Two-view Auto-stereoscopic Display	30
2.14	Multi-view displays create multiple viewing windows	32
2.15	Tracking Two-view Displays	34
2.16	Illustration of a natural 3D display	35
2.17	Head-mounted Display [HMD, 2011]	40
3.1	Vitaphone System [Vit, 2011]	44

3.2	Cinerama Diagram	46
3.3	Anamorphic Lenses	46
3.4	DLP Projectors	46
3.5	Stereoscopic Film Projection	47
3.6	3D Digital Projection System	48
3.7	Titmus Fly Stereotest	50
3.8	Teleview System [Tel, 2011]	51
3.9	3D Eyewears	52
3.10	The viewing frustum	55
3.11	Converged camera model (left) and parallel camera model (right) [Holliman et al., 2006]	55
3.12	Vertical Disparity [Woods et al., 1993]	56
3.13	Symmetric parallel camera configuration [Froner and Holliman, 2005]	57
3.14	A Floating Window	59
3.15	The Multi-rigging Configuration	61
3.16	Perceived depth change caused by object movement	67
3.17	Perceived Depth change caused by camera motion	68
4.1	Williams and Parrish's results on the limits of the 3D Comfort Zone	72
4.2	An example of a static scene and dynamic scene	76
4.3	The cyclopean scale	77
4.4	An illustration of the distortion caused by false-eye modelling	78
4.5	Jones <i>et al</i> 's and Holliman's Algorithms	80
4.6	Jones <i>et al</i> 's algorithm in a dynamic stereoscopic scene	81
5.1	Real-Eye Configuration	89
5.2	Jones <i>et al</i> 's algorithm	90
5.3	Jones <i>et al</i> 's fixed mapping method vs. our dynamic mapping method	94
5.4	Fixed DOF blur simulation	99
5.5	Dynamic DOF blur simulation	100
5.6	Experiment Procedure	102
5.7	Experiment Apparatus	103

5.8	Scoring scale	104
5.9	Real-Eye Stimuli	105
5.10	Fixed Mapping Stimuli	106
5.11	Dynamic Mapping Stimuli	106
5.12	Fixed DOF Stimuli	106
5.13	Dynamic DOF Stimuli	107
5.14	An example of a result sheet	108
5.15	Results of each method tested in the experiment	109
5.16	Box Plot Results	110
6.1	Random Dot Stereogram	115
6.2	Random Dot Autostereogram [aut, 2011]	116
6.3	Menu for selecting the shape	123
6.4	Experiment Apparatus	124
6.5	Stimulus in the pre-test session	125
6.6	Stimulus in the test session	125
6.7	Results	126
6.8	Mistakes made on the depth planes in front of the display	128
6.9	Mistakes made on the depth planes behind the display	130
6.10	Time spent on the depth planes in front of the display	131
6.11	Time spent on the depth planes behind the display	132
6.12	Mistakes made for square and triangle in front of the display	134
6.13	Mistakes made for square and triangle behind the display	135
6.14	Time spent on square and triangle in front of the display	136
6.15	Time spent on square and triangle behind the display	137
7.1	Fixed Depth Mapping vs. Dynamic Depth Mapping	143
7.2	Depth of Field Blur Simulation	144

List of Tables

3.1	Innovations in Cinema Technology [Lipton, 2007]	43
3.2	Cinematic Storytelling Techniques	63
5.1	Means and Standard Deviations	111
5.2	T-Test results concerning the Real-Eye method	111
5.3	T-Test results between mapping methods and DOF simulations	112
6.1	Successes made on the depth planes in front of the screen	129
6.2	Time spent on the depth planes in front of the screen	131
6.3	Time spent on the depth planes behind the screen	133
6.4	Mistakes made for square and triangle in front of the screen	134
6.5	Time spent on square and triangle in front of the screen	137
6.6	Time spent on square and triangle behind the screen	139
7.1	Results on the limits of the 3D Comfort Zone	141

List of Algorithms

1	Initiating the Frame Buffer Object	95
2	Updating Scene Boundaries with Z-Buffer	97
3	Off-screen Rendering	98
4	Drawing RDS Image	118
5	Rendering Stereoscopic RDS image with Stencil Buffer	119

Chapter 1

Introduction

This chapter describes the basic background, the research problem, research methods and contributions, and the organisation of this thesis.

1.1 The Renaissance of the Stereoscopic Cinema

The stereoscopic (S3D) cinema was once at its peak in the 1950s. More than 65 stereoscopic feature films were produced by the Hollywood from the year 1952 to 1954 [Lambooi et al., 2007]. However, the popularity of the medium rapidly declined due to the technical problems such as the film projection glitch, images misalignment, and visual discomfort [Lambooji et al., 2009]. In 2005, the single digital S3D projection system was developed by RealD (Beverly Hills, CA, USA), which solved most of those technical problems and triggered the recent explosion of producing and projecting movies in the S3D format [Lipton, 2007], as illustrated in Figure 1.1.

Quite a few stereoscopic movies with decent quality have been released over the past few years, such as Intel's *CyberWorld 3D* in 2000, *Santa vs the Snowman* from Universal Studio in 2002, *The Polar Express* (2004) and *Beowulf* (2007) from Sony Pictures and *Avatar* (2009) from Twentieth Century Fox. More than 100 mainstream S3D movies, including blockbusters such as *Kung Fu Panda 2* from DreamWorks, *Toy Story 3* from Pixal Animation Studios, and *Harry Potter and the Deathly Hallows: Part 1 & 2* from Warner Bros. Pictures, were released in the last two years (2010 and 2011) both in RealD and IMAX 3D cinemas worldwide [3dm,

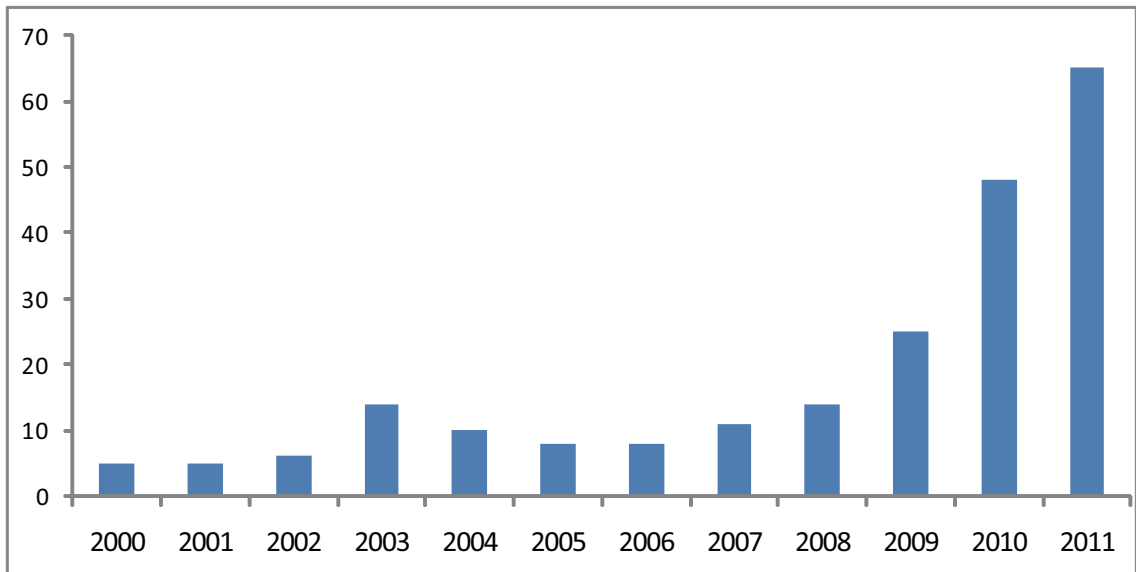


Figure 1.1: Numbers of 3D movies released since the year 2000 [3dm, 2011]

2011]. The renaissance of the S3D cinema is upon us [Autodesk, 2008, Lang et al., 2010, Koppal et al., 2010, Lipton, 2007].

However, the stereoscopic cinema’s resurgence in popularity has been accompanied by serious concerns about adverse effects regarding the S3D viewing discomfort [Shibata et al., 2011]. The practical construction of stereoscopic content that leads to a comfortable viewing experience remains to be a great research challenge due to the complexity of the human visual system and the restrictions of stereoscopic displays [Howard and Rogers, 2002, Hoffman et al., 2008, Lang et al., 2010].

1.2 The Research Problem

This thesis aims to reduce the S3D viewing discomfort, in particular, the viewing discomfort caused by the so-called “**Vergence-Accommodation (VA) Conflict**”.

1.2.1 The Vergence-Accommodation Conflict

When viewing stereoscopic displays the eye vergence and accommodation conflict with each other constantly: the eyes must maintain focus (accommodation) on the display screen but at the same time rotate (verge) away from the display screen to where the 3D object is perceived. This effect is illustrated in Figure 1.2. The

difference between where eyes are verged and accommodated creates the VA Conflict. The degree of the VA Conflict is decided by the magnitude of the perceived depth, p , which is the distance between the perceived 3D object and the display screen.

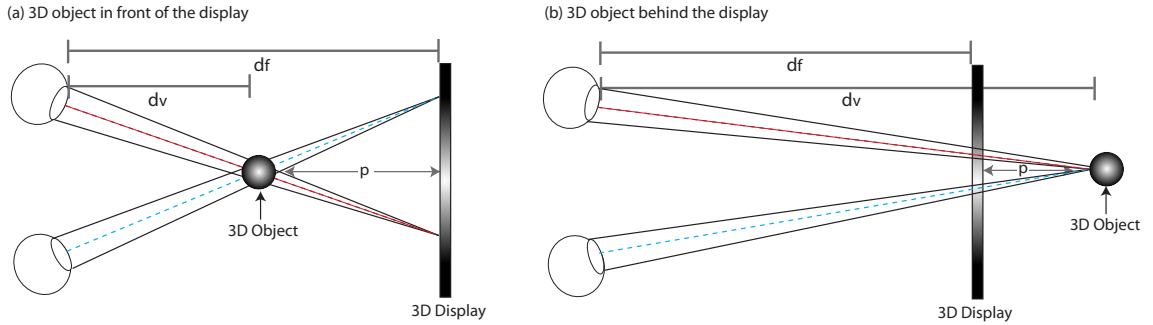


Figure 1.2: Vergence-Accommodation Conflict in S3D viewing. d_f is the focus distance, d_v is the vergence distance, p is the perceived depth.

Consider Figure 1.2(a), the focus distance (d_f) is larger than the vergence distance (d_v) ($d_f > d_v$) when the 3D object is perceived in front of the screen; In Figure 1.2(b), the focus distance is smaller than the vergence distance ($d_f < d_v$) when the 3D object is perceived behind the 3D screen. Note that when the object is perceived on the screen, the focus distance is equal to the vergence distance ($d_f = d_v$). There is no conflict between eye vergence and accommodation, just as watching an object on a 2D screen.

We focused on the VA Conflict for two reasons:

- The VA Conflict is present in all types of stereoscopic display systems (3D mobiles, 3D desktop displays, 3DTVs, 3D cinemas, etc.) [Shibata et al., 2011].
- All human factors studies agreed that the VA Conflict is a great contributor to the visual discomfort when viewing S3D content [Yeh and Silverstein, 1990, Hiruma and Fukuda, 1993, Wopking, 1995, Wann and Mon-Williams, 1997, Sugihara et al., 1999, Jones et al., 2001, Howard and Rogers, 2002, Yano et al., 2002, Wann and Mon-Williams, 2002, Yano et al., 2004, Emoto et al., 2005, Häkkinen et al., 2006, Hoffman et al., 2008, Ukai and Howarth, 2008, Lambooji et al., 2009, Shibata et al., 2011].

1.2.2 3D Comfort Zone

Studies have confirmed that the level of S3D viewing discomfort is in proportion with the degree of the VA conflict [Sugihara et al., 1999, Hoffman et al., 2008]. This phenomenon creates a so-called “**3D Comfort Zone**” around the 3D display screen [Mendiburu, 2009].

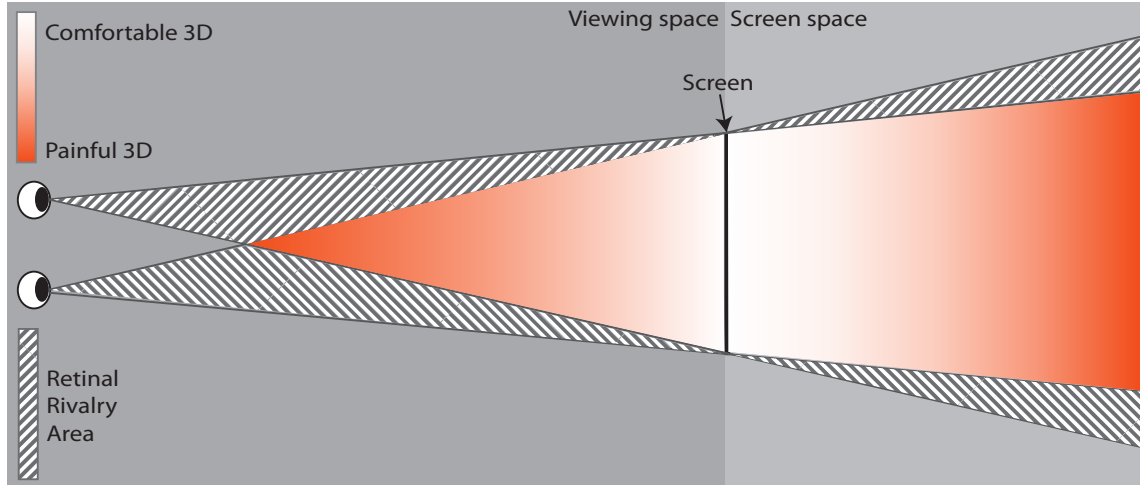


Figure 1.3: An illustration of the 3D Comfort Zone [Mendiburu, 2009]. The white-to-orange gradient illustrates the stereoscopic viewing area. The 3D Comfort Zone is from white to light orange.

Consider Figure 1.3, the 3D Comfort Zone, from white to light orange, is the region where the conflict between eye vergence and accommodation is small. The perceived depth inside the 3D Comfort Zone does not cause S3D viewing discomfort [Ronfard and Taubin, 2010]. As the conflict between eye vergence and accommodation increases, the perceived depth, from light orange to dark orange, exceeds the 3D Comfort Zone and becomes more and more painful for viewers. This is the reason that human factors studies [McAllister, 1993, Shibata et al., 2011] recommend limiting the perceived depth inside the 3D Comfort Zone for comfortable S3D viewing.

Note that the striped patterns in Figure 1.3 are Retinal Rivalry Areas which are seen by only one eye of the viewer. They should be used with caution [Mendiburu, 2009].

1.3 Research Methods and Contributions

In order to reduce the S3D viewing discomfort imposed by the VA Conflict, we studied the following three topics: limits of the 3D Comfort Zone, depth mapping methods, and the Depth of Field (DoF) blur technique.

1.3.1 Limits of the 3D Comfort Zone

As discussed above, the perceived depth must be limited inside the 3D Comfort Zone on the 3D display in order to ensure a comfortable S3D viewing experience. Consequently, identifying the limits of the 3D Comfort Zone is a prerequisite to effectively controlling the perceived depth in S3D.

Research has confirmed that the limits of the 3D Comfort Zone vary among 3D display systems due to different display characteristics [Holliman et al., 2007] and different viewer-to-display (viewing) distances [Shibata et al., 2011]. We proposed a new method that employed the **Random Dot Stereogram** (RDS) technique to efficiently identify the practical limits of the 3D Comfort Zone on a given 3D display.

1.3.2 Depth Mapping Methods

Depth mapping in S3D refers to mapping the scene depth into a predefined perceived depth range on the 3D display, as illustrated in Figure 1.4. Note that the Zero Disparity Plane (ZDP, also known as the virtual display) is a plane in the scene. The ZDP is to be mapped onto the physical display so that scene objects that are in front of the ZDP are seen in front of the physical display; scene objects behind the ZDP are seen behind the physical display; and scene objects on the ZDP are seen on the physical display. Scene depth is the distance between the 3D object and the ZDP in the scene.

Existing depth mapping methods can limit the perceived depth inside the 3D Comfort Zone to avoid excessive perceived depth. However, they suffer major drawbacks associated with inadequate perceived depth, and dramatic and sudden change of the perceived depth. In this study, we developed a new depth mapping method that dynamically adjusts the mapping from scene depth to perceived depth. It

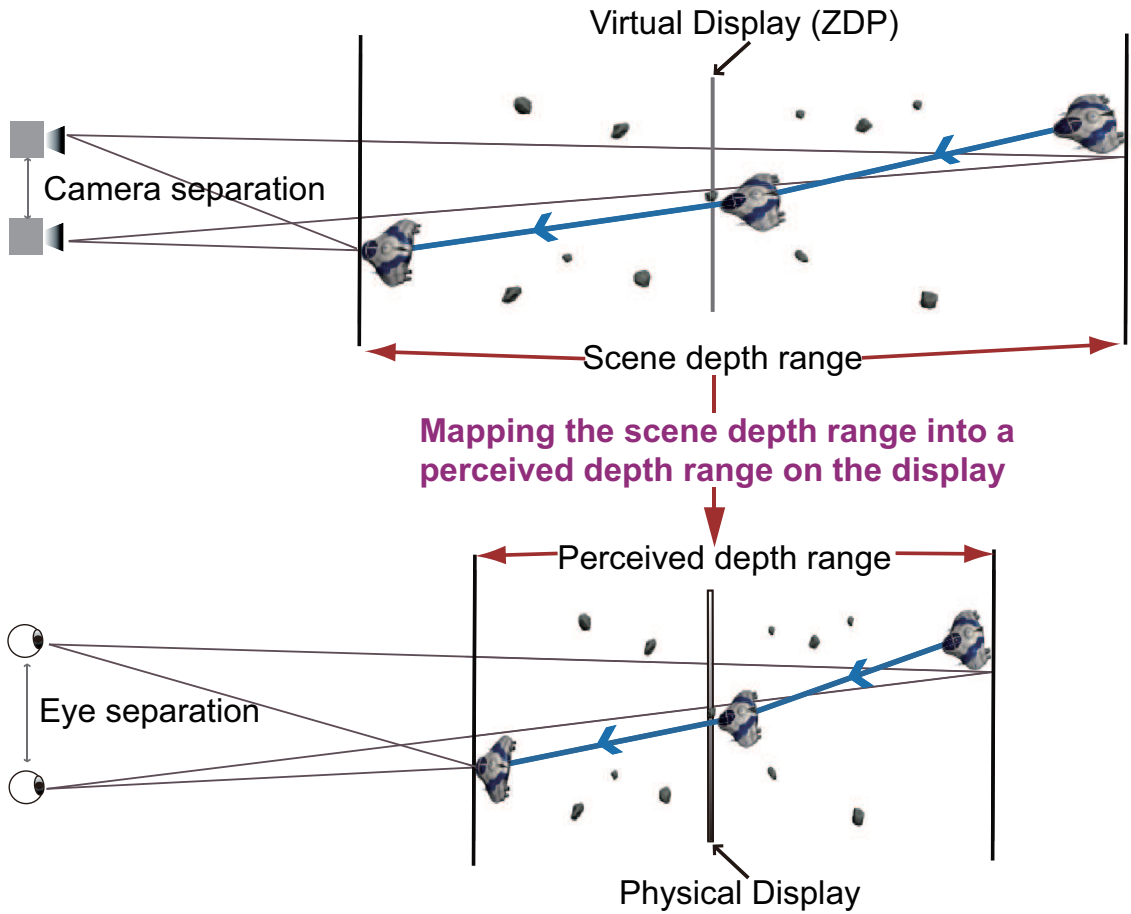


Figure 1.4: Mapping of the scene depth range (top) to a perceived depth range (bottom).

ensures that the viewer's perceived depth stays constant on the 3D display and is always mapped to the limits of the 3D Comfort Zone, providing maximum perceived depth without causing the viewing discomfort. The results of a human-based experiment confirmed that our method did provide better S3D viewing experience than existing depth mapping methods.

1.3.3 Depth of Field Blur Technique

The Depth of Field (DoF) refers to a range inside which objects are seen clear and sharp, and outside which objects are seen blurred. By creating a DoF around the 3D display, objects close to the 3D screen are placed inside the DoF. Therefore, they are seen in full sharpness. Objects that are far away from the 3D screen and create

perceived depth exceeding the limits of the 3D Comfort Zone, are placed outside the DoF and seen blurred. As viewers can not focus on blurred objects, there is no VA Conflict. This effect is illustrated in Figure 1.5.

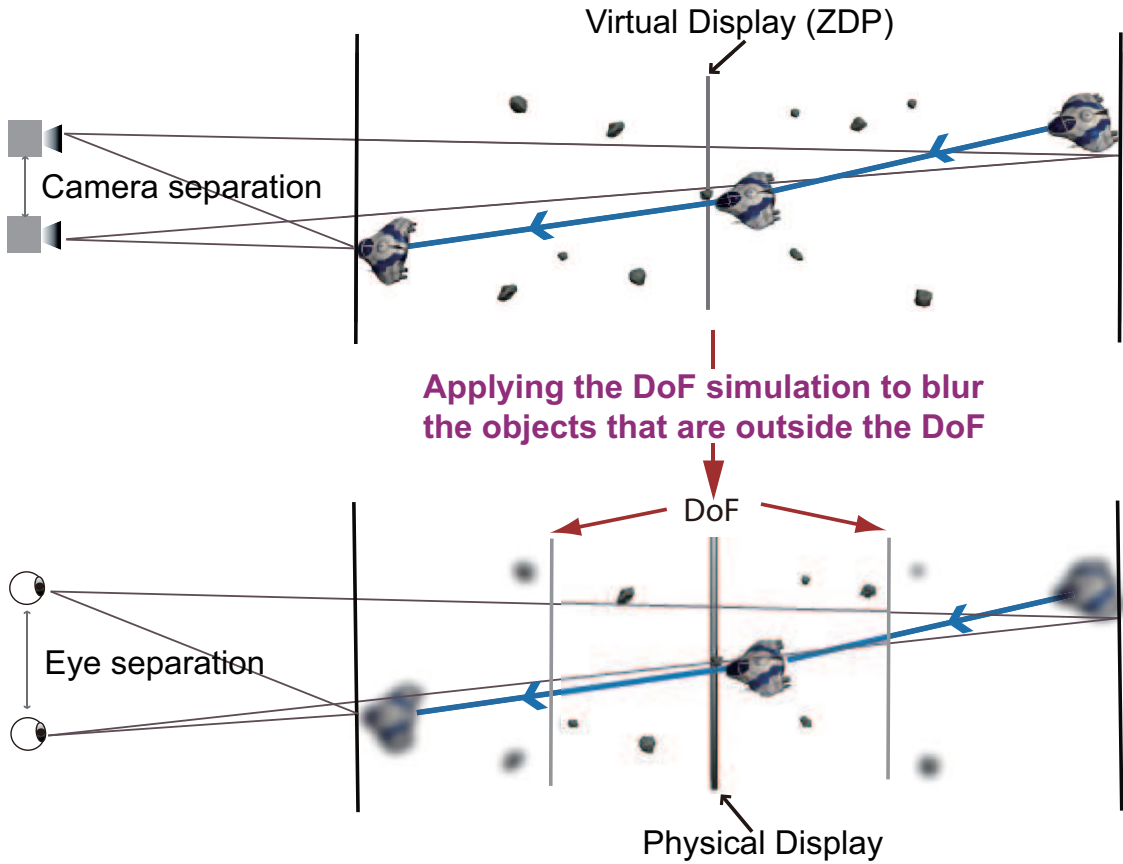


Figure 1.5: The DoF blur simulation. Objects inside the DoF are seen in full sharpness; objects outside the DoF are blurred.

We implemented two DoF simulation approaches (one with a deep and static DoF and another with a shallow and dynamic DoF), and evaluated them in a human-based experiment. Our results suggested that viewers may not enjoy watching stereoscopic content with the DoF simulation. The DoF should be used with caution.

1.4 Thesis Organisation

Chapter 2 presents general background information concerning the framework for understanding stereoscopic displays and applications. The chapter reviews human

binocular visual system and the development and applications in modern 3D display systems, ranging from the traditional stereoscopic displays to the most state of the art natural 3D displays. Technologies such as volumetric, holographic, and head-mounted stereoscopic displays are also briefly discussed.

Chapter 3 reviews the history of the cinema technology from the invention of sound and colour to the stereoscopic cinema, explores the reasons for the recent renaissance of the stereoscopic cinema, analyses technical issues in S3D, discusses research problems brought by applying *Cinematic Storytelling* techniques in stereoscopic cinematography, and finally, summarises well-known key factors in producing a comfortable viewing S3D movie and discusses remaining research challenges in stereoscopic cinematography.

Chapter 4 reviews the literature in efficiently controlling the perceived depth in S3D. Detailed description of how previous studies have been carried out is provided. The limitations of existing works are summarised along with the motivation for conducting new research on stereoscopic depth perception control in S3D.

Chapter 5 investigates five different perceived depth control approaches: a new dynamic depth mapping method, an existing fixed depth mapping method, two DoF blur simulation approaches, and a real-eye configuration method. The next section of this chapter provides a detailed description of the human-based trial performed to evaluate the above-mentioned stereoscopic imaging approaches. This section starts with an introduction of the experiment protocol, including the objective and hypothesis of the experiment, the apparatus and viewing conditions, and the procedure of the experiment, followed by describing the stimulus used in the trial and how we implemented them. Finally the statistical analysis of the experimental results are discussed.

Chapter 6 is dedicated to identifying the limits of the 3D Comfort Zone on a given desktop 3D display. A novel method that uses the RDS technique to identify the range of the 3D Comfort Zone is provided. The details of the experiment conducted to evaluate this method and the statistical analysis of the experiment results are also presented.

Chapter 7 compares our results with previous studies, discusses our findings in

terms of their meaning in stereoscopic cinematography, makes our own recommendations on how to better control depth perception in stereoscopic cinematography, discusses the limitations of this work, and finally, establishes a basic guideline for stereoscopic cinematography, following which can help ensure a compelling and comfortable 3D viewing experience.

Finally, Chapter 8 concludes the thesis and discusses the future work.

Chapter 2

Background

2.1 Human Visual System

Binocular vision provides humans with the advantage of depth perception. The depth perception is based on the discrepancies in the location of corresponding points in the left and right images incident on the eyes, which the brain interprets as depth. This is known as Stereopsis. Human visual system also uses other depth cues to perceive depth. We consider them later in this section.

2.1.1 The Human Eye

Figure 2.1 illustrates the optical model of the human eye and its components. The human eye works in a similar way to a camera. The **Cornea** is the transparent front part of the eye and acts as the lens cover of a camera. It allows widely diverged light rays to pass through the **Iris** which behaves much like the aperture in a camera, adjusting the eye to different lighting conditions. Next in the line is the **Lens** which acts very similarly as the lens of a camera, focusing rays of light to the “film” of the eye: **Retina**. As illustrated in Figure 2.1, the retina is a sub-portion of the interior surface of the eye, starting at the **Ora Serrata**. The retina contains photoreceptors and associated neural processing circuitry which convert light rays to impulses and send them via optic nerves to the brain, where an image of what the viewer sees is perceived.

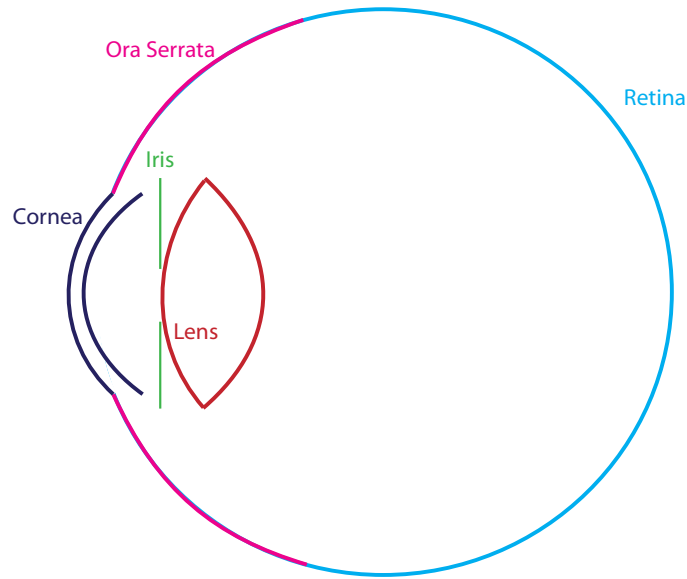


Figure 2.1: Optical model of the human eye [Deering, 2005]

2.1.2 Binocular Depth Perception in the Natural World

In order to understand binocular depth perception in electronic images (computer graphics), one must start with a basic understanding of the binocular depth perception in the natural world.

When human eyes “look” at an object in the real world, the eyes rotate to look toward the fixation point and adjust the shape of Lens in order to bring this fixation point into focus. The rotation process of eyes is called **Vergence**. Vergence onto a new fixation point may require *Convergence* or *Divergence* depending on the spatial locations of the previous and new fixation points. *Convergence* occurs if the new fixation point is closer (to viewer) than the previous fixation point. The eyes must rotate inwardly relative to their previous orientations. *Divergence* is rotating the eyes outwardly relative to their previous orientations when the new fixation point is further away from the viewer. In addition to the Vergence, the adjusting of Lens’s shape is called **Accommodation**. Accommodation guarantees that the fixation point lies on the retina’s highest resolution portion, Fovea.

Panum's Fusion

Figure 2.2 illustrates the area of the **Panum's Fusion**. Points **A** and **F** lie inside the Panum's Fusion and are seen as single images. However, point **B** falls outside the Panum's Fusion and cannot be fused by the human visual system. Point **F** is on the **Horopter** where points are projected to the same location on each retina therefore **F** has zero disparity between its location in the left and right retinal images. The shape of the horopter shown in Figure 2.2 is for illustration only. Research suggests that the Horopter's shape is complex and nonlinear [Blakemore, 1970, Goldstein, 2002].

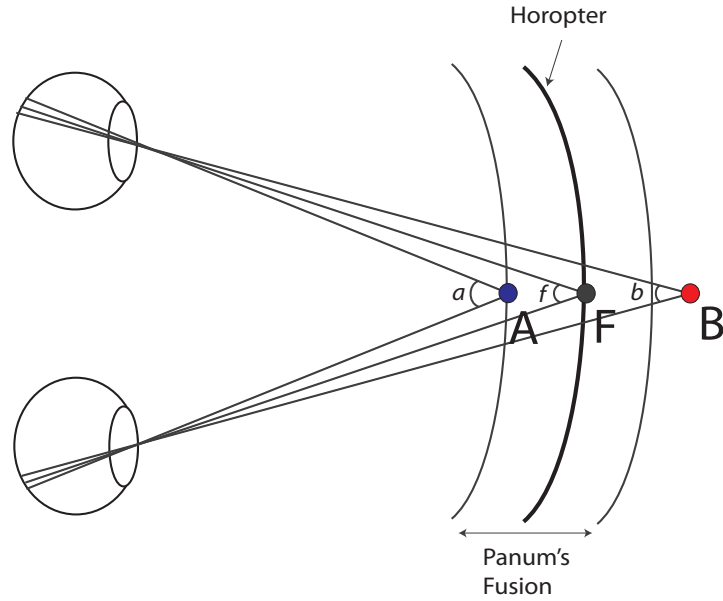


Figure 2.2: Illustration of the Panum's Fusion

Angular Disparity

Consider Figure 2.2, **Angular Disparity**, AD , is defined as the difference between the vergence angle at the fixation point **F** and the point of interest.

Point **A** lies in front of **F**, it has negative angular disparity:

$$AD = f - a \quad (AD < 0) \quad (2.1)$$

Point **B** lies behind **F**, it has positive angular disparity:

$$AD = f - b \quad (AD > 0) \quad (2.2)$$

Objects with negative angular disparities are seen in front of the focus while objects with positive disparities are seen behind the focus.

Stereo Acuity

Stereo Acuity is defined as the smallest perceptible binocular depth difference between two objects. Defining stereo acuity as an angle provides the advantage that it can be assumed to be constant regardless of the actual distance to and between the two points. It is also useful to know how this angle translates to the smallest perceived distance between objects because it will allow us to compare the eye's ability of perceiving depth with the display's ability of reproducing it [Holliman, 2006].

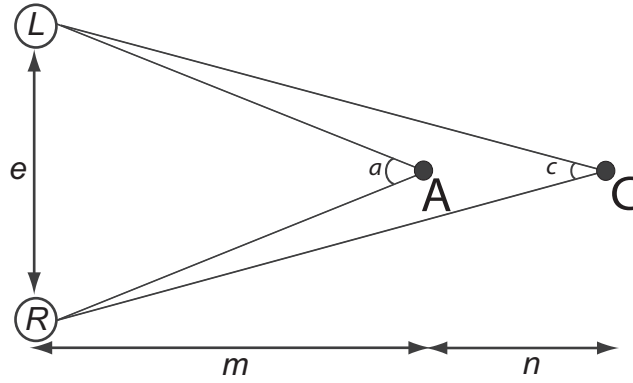


Figure 2.3: Stereo acuity defines the smallest perceptible depth difference

Considering Figure 2.3, assuming points A and C can be just distinguished at different depth planes, the stereo acuity, δ , will be:

$$\delta = a - c \quad (2.3)$$

Studies [Julesz, 1971, Langlands, 1926, Yeh and Silverstein, 1990] show that human eyes can distinguish as little as $1.8''$ (seconds of arc) of δ . Although the exact

limit varies among subjects, Diner and Fender [Diner and Fender, 1993] suggested that a practical working limit for the stereo acuity is a value of $\delta = 20''$. Using this value we can calculate the distance of the smallest distinguishable depth difference at a given distance to the viewer.

Taking $m = 750mm$ as the common viewing distance from the viewer to typical desktop stereoscopic displays and eye separation, $e = 65mm$ [Ronfard and Taubin, 2010], the minimum distinguishable depth, n , at the distance m can be calculated by considering points **A** and **C**. The angle a can be derived from:

$$a = 2 * \arctan \left(\frac{(e/2)}{m} \right) = 2 * \arctan \left(\frac{65/2}{750} \right) \quad (2.4)$$

by the definition of stereo acuity we have:

$$\tan(c/2) = \tan \left(\frac{a - \delta}{2} \right) = \tan \left(\frac{a - 20''}{2} \right) \quad (2.5)$$

and as n is the distance between **A** and **C** we know that:

$$\tan(c/2) = \frac{(e/2)}{m + n} \quad (2.6)$$

rearranging (2.6) we have:

$$n = \left(\frac{e/2}{\tan(c/2)} \right) - m \quad (2.7)$$

substituting (2.4) in (2.5) and using the result in (2.7), we have $n = 0.84mm$.

We conclude that the smallest perceptible depth differences is 0.84mm for a person with a stereo acuity of $20''$, eye separation of 65mm and viewing distance of 750mm.

It is also useful to calculate the value of the furthest limits of binocular depth perception which can be reached when the two visual axes' vergence angle, a , is equal to or less than the angle of stereo acuity.

Rearranging (2.4) we know that:

$$m = \frac{e/2}{\tan(a/2)} \quad (2.8)$$

Again choosing $\delta = 20''$ and $e = 65mm$, when the vergence angle $a = \delta$, we have $m = 670mm$.

This indicates that points such as **C** at the distance of $670m$ or more from the viewer cannot be perceived at a different depth plane as **A** by using binocular vision alone [Holliman, 2006].

2.1.3 Binocular Depth Perception in Electronic Stereoscopic Images

The stereoscopic depth sensation was first recreated by Wheatstone [Wheatstone, 1838] in 1838. Two planar views of a 2D planar image of the same content from slightly different viewpoints were shown to left and right eyes as illustrated in Figure 2.4. The viewpoint difference generated retinal image disparity which created perceived binocular depth when the observer subsequently viewed the images. Wheatstone built the first stereoscope to demonstrate this effect and many other devices have since been invented for presenting stereoscopic images. Reviews on these devices are available in [Helmholtz, 1962, Valyus, 1966, Lane, 1982, Lipton, 1982, Okoshi, 1976].

Geometry of Perceived Binocular Depth

The geometric model of perceived depth, GPD, has been studied by Helmholtz [Helmholtz, 1962] and Valyus [Valyus, 1966] and more recently in [Hodges and Davis, 1993, Woods et al., 1993, Diner and Fender, 1993, Jones et al., 2001]. In this study, we analyse a simplified geometric model illustrated in Figure 2.5 which emphasises on the key variables affecting the binocular depth perception of electronic stereoscopic images.

The observer's eyes, **L** and **R**, are separated by the interocular distance, e , and are at a viewing distance, z , from the stereoscopic display. p is the perceived depth on the display. d , the screen disparity between two corresponding points in the left and right images, is a physical distance on the screen resulted by the image disparity. Image disparity is a logical value measured in pixels and it is constant for a given stereo pair while the screen disparity varies depending on the characteristics (size, ratio, etc) of the display.

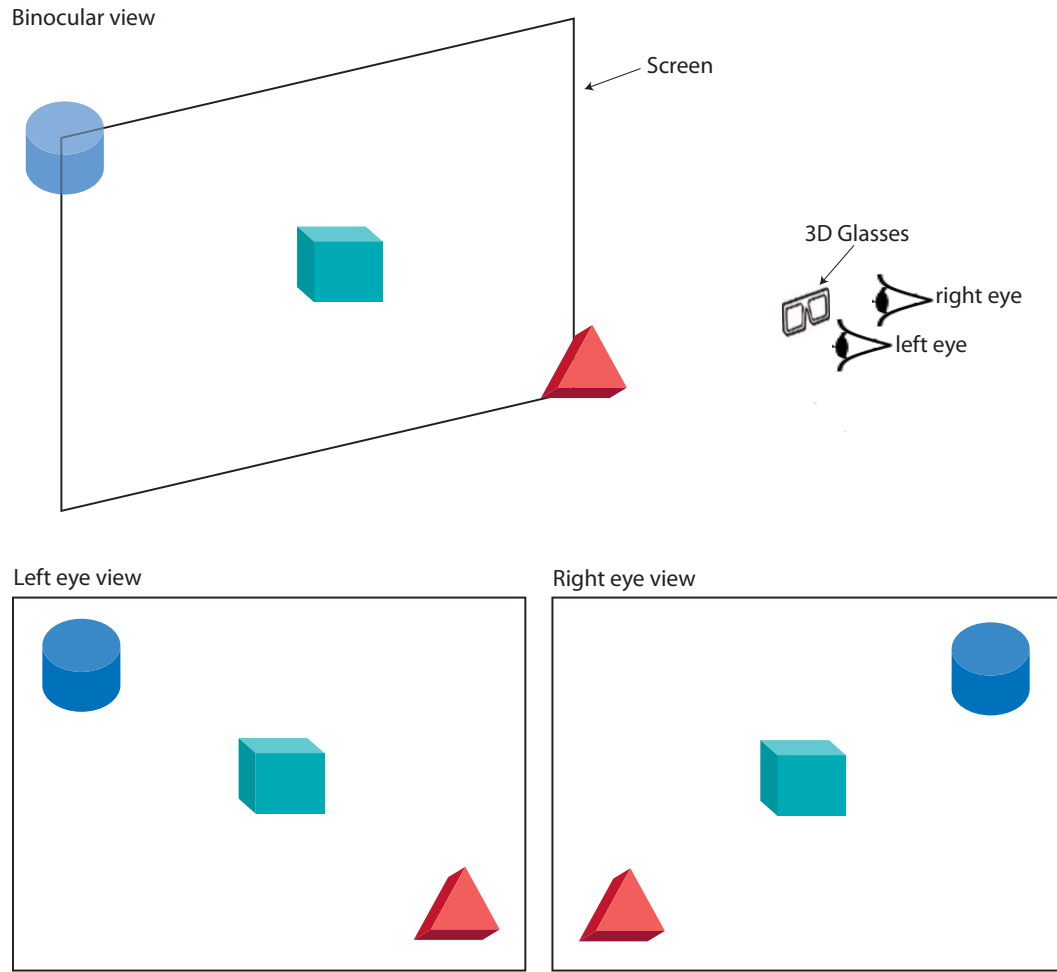


Figure 2.4: Binocular view in electronic stereoscopic images

Screen Disparity

As shown in Figure 2.5, the screen disparity in a stereoscopic pair is the difference of the physical x coordinates between corresponding points in the right x_r and left x_l images.

$$d = x_r - x_l \quad (2.9)$$

Positive Disparity In the top part of the Figure 2.5, the screen disparity, d , has a positive value which is also referred to as the uncrossed disparity. The eyes rotate outwardly in order to re-fixate on a point further behind the screen. The uncrossed disparity introduces the stereoscopic depth behind the screen.

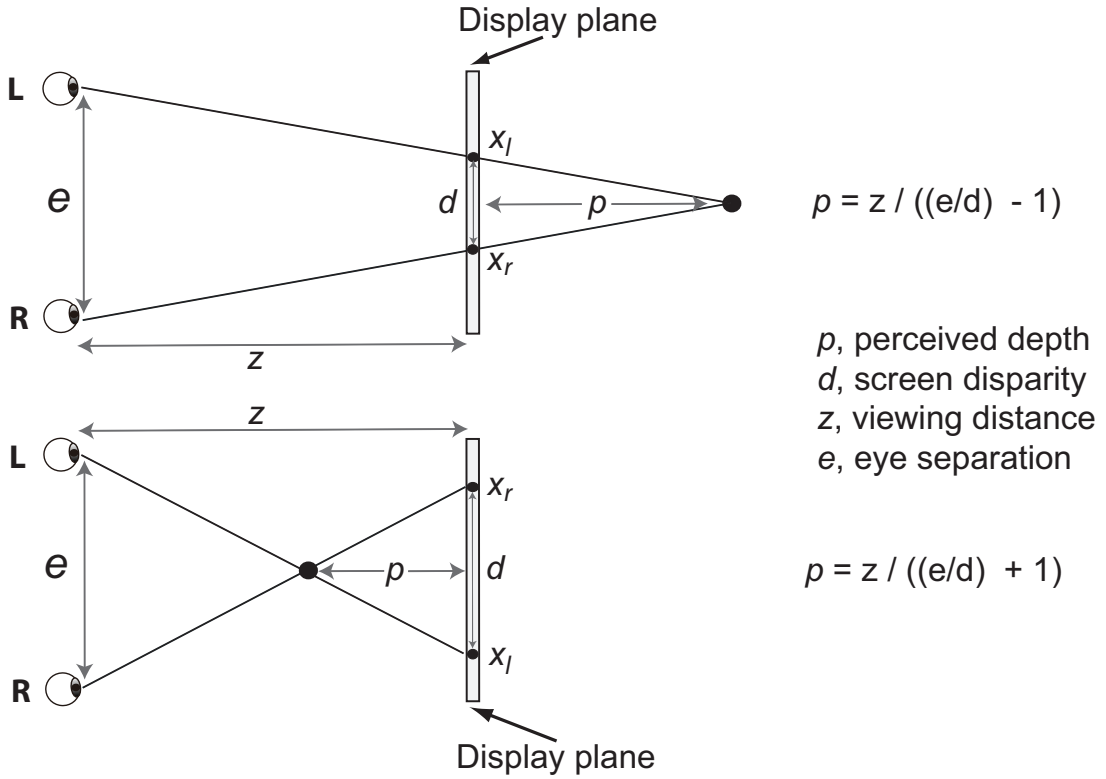


Figure 2.5: Perceived depth behind (top) and in front (bottom) of the display

The perceived depth with uncrossed disparity is given by:

$$p = \frac{z}{\frac{e}{d} - 1} \quad (2.10)$$

If a point is on the display screen, then the screen disparity, d , is equal to zero. As a result, the perceived depth, p , is also equal to zero. Note that when the positive disparity is equal to the interocular distance, the 3D object appears to be at infinity. This particular positive disparity is called ***Infinite Disparity***.

Negative Disparity Consider the right part of the Figure 2.5, the screen disparity, d , has a negative value. The negative screen disparity can also be referred to as the crossed disparity. This is because if the eyes are previously fixated on the projection plane, their visual axes must cross each other so that the eyes can re-fixate on a point in front of the screen. The crossed disparity introduces perceived stereoscopic depth in front of the screen.

The perceived depth with crossed disparity can be derived from the similar tri-

angles in Figure 2.5:

$$p = \frac{z}{\frac{e}{|d|} + 1} \quad (2.11)$$

Together with Positive and Negative Disparities, Figure 2.6 demonstrates two other types of screen disparities:

- Figure 2.6(A) illustrates the ***Zero Disparity***. When the left and right stereo pairs directly lie on top of each other, the screen disparity is zero and there is no difference between left and right images.
- Figure 2.6(D) shows the ***Divergent Disparity***. When the screen disparity is actually larger than the human eye's interocular distance, the viewer's lines of sight has to be diverged in order to perceive the left and right images at the same time. The Divergent Disparity never takes place in the real world and it is strongly recommended that the Divergent Disparity be avoided in S3D content [Spottiswoode et al., 1952, Lipton, 1982].

Several geometric factors affecting the binocular depth perception can be drawn from equations (2.11) and (2.10):

- z , the viewing distance from viewer to the display. The perceived binocular depth, p , is proportional to the viewing distance, z . Therefore increasing viewing distance provides increased perceived binocular depth and *vice-versa*.
- d , the screen disparity. The perceived binocular depth, p , is also directly proportional to the screen disparity, d . Increasing the size of the stereoscopic image by either increasing the image size on the same display plane or displaying the same image at a larger screen provides more perceived binocular depth and *vice-versa*.
- e , viewer's eye separation. The perceived binocular depth, p , is inversely proportional to the viewer's eye separation, e . The viewer's eye separation varies between 55mm to 75mm approximately. The average value is often taken as 65mm. Generally, children have smaller values of eye separation and therefore perceive more binocular depth in the same stereoscopic image than adults [Dodgson, 2004].

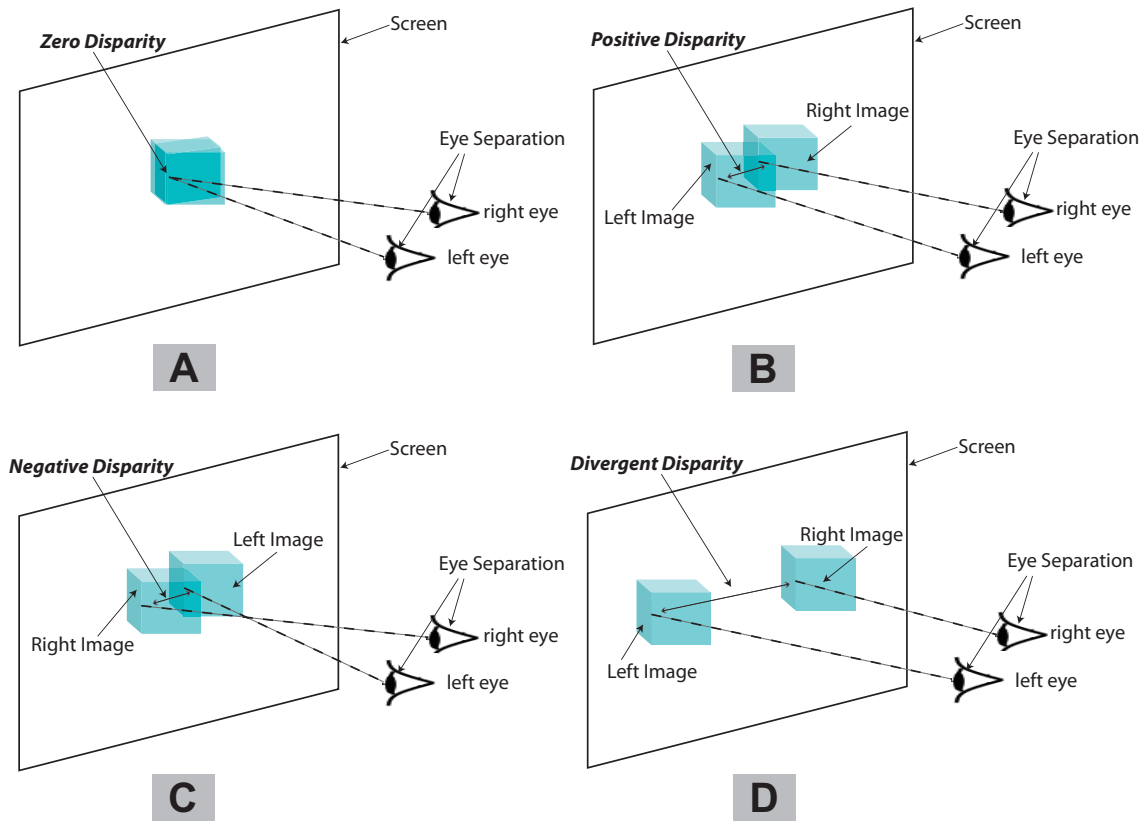


Figure 2.6: Four types of screen disparities. (A) The object has Zero Disparity appearing on the display screen. (B) Positive Disparity refers to the screen disparity that is larger than zero, but smaller than/equal to the interocular distance. (C) Negative Disparity occurs when the axes of the eyes cross in front of the display screen. (D) When the screen disparity is larger than the interocular distance, the Divergent Disparity occurs.

2.1.4 Other Depth Cues

The human visual system also makes use of monocular depth cues (also known as pictorial [Goldstein, 2002] or empirical cues [Ogle, 1964]), oculomotor cues, and motion parallax to help determine relative depth in the 3D world.

Monocular Depth Cues

Research shows that even people with monocular vision can perform good depth judgement in the real world [Bruce et al., 2003]. Therefore, it is important to understand the contribution of monocular 2D depth cues for designing stereoscopic

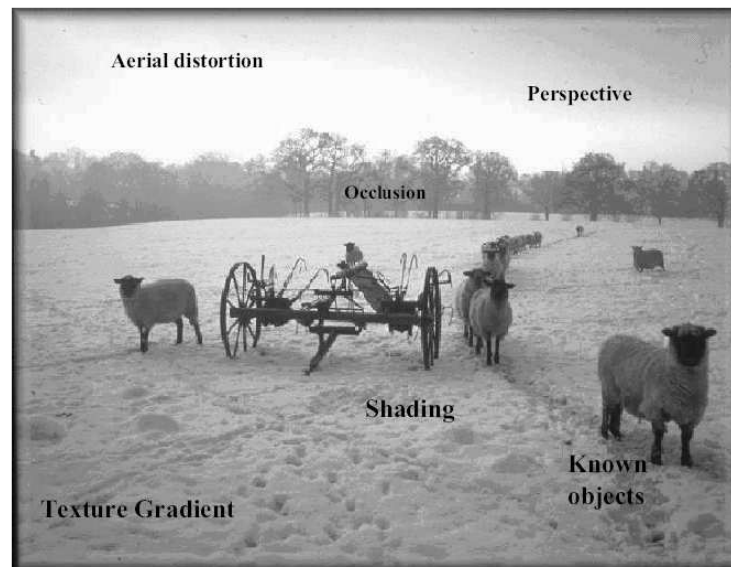


Figure 2.7: Monocular depth cues available in a 2D image (photographer David Burder)

displays. Ezra *et al* [Ezra et al., 1995] suggested that 3D displays should have comparable brightness, contrast, resolution, and viewing range as standard 2D displays. The available monocular depth cues are demonstrated in Figure 2.7, which are:

Linear perspective The image size of a target object on the retina is inversely proportional to the distance between the object and the observer. Hence, the retinal image size becomes smaller as an object moves further away and *vice-versa*.

Shading and shadowing The amount of reflected light on a surface changes in inverse proportion to the distance from the light source to the surface. Surfaces that are further away from the light source appear to be darker than closer ones which provides cues to their depth relationships. Shadows casted by one object on another also give salient depth cues to their relative positions.

Aerial Perspective Objects that are closer to the observer appear to be more distinct than those that are further away as atmosphere affect the light travelling. Distant objects tend to be bluish because the colour blue has a shorter wavelength and can penetrate the atmosphere more easily than other colours.

Interposition Objects occluding or overlapping each other suggests their relative positions. The object that is creating the occlusion appears to be closer to the viewer.

Retinal Image Size An object with larger retinal image size is judged closer to the viewer than the same object with a smaller retinal image size and *vice-versa*.

Texture Gradient More details can be perceived in objects that are closer to the viewer. As objects move further away, the texture becomes blurred.

Colour Human eyes refract different colours at different angles. Therefore objects with the same shape and size and at the same distance are perceived at different depths because of colour differences. In general, bright-coloured objects tend to appear closer to the viewer than dark-coloured ones.

Motion Parallax and Oculomotor Depth Cues

Two other non-binocular depth cues are available:

Motion Parallax provides different views of a scene due to the viewer or scene movement. Consider the situation in [Ware and Franck, 1996], a cloud of discrete nodes in a 3D space in which all of them have the same colour and size. The depth judgement can be made by noticing how much two nodes move relative to each other when the viewer moves side to side or up and down: nodes that are closer to the viewer appear to move more than nodes that are further away. Motion parallax provides a powerful cue for judging relative depth in a 3D scene without the use of stereopsis [Ogle, 1964, Goldstein, 2002]. However, studies have shown that motion parallax does not make the stereopsis redundant; combining stereopsis and motion parallax results in better depth perception than using either cue alone [Yeh, 1993].

Oculomotor Depth Cues are due to the muscle movements during eyes' vergence and accommodation. They are regarded as having little potential in helping the depth judgement [Glassner, 1995, Ogle, 1964, Pastoor, 1991].

2.1.5 Benefits of the Binocular Vision

An important question is what benefits can the binocular vision actually provide over monocular vision in the real life? The majority of people regard seeing an image in three-dimensional on a planar surface as a fascinating visual effect. Beyond the obvious attractiveness of perceiving images in 3D, binocular vision also have the following advantages over monocular vision:

- Judgement of relative depth. It is difficult to judge the relative depth between objects in space using only monocular depth cues, especially when objects are distant from viewers. Binocular vision can significantly improve the effectiveness of the relative depth judgement.
- Localisation in space. Binocular vision helps the brain focus on objects at a certain depth plane and ignore other objects placed at other depth planes.
- Surface curvature interpretation. There is evidence suggesting that binocular vision can offer better interpretation of the curved surfaces [Holliman, 2004].
- Perception of surface material. Certain lustrous substances, such as glittering metals and sparkling gems, are perceived as such partly because of the different specular reflections detected at the same time by both the left and right eyes [Helmholtz, 1962].
- Identifying camouflage. According to [Schiffman, 2000], one of the most fundamental reasons to have binocular vision is to identify objects with camouflage in a scene.

Certain professional (medical, military, etc) applications, in which the judgement of depth is vital to the successful accomplishment of tasks, have been considerably benefited by the display and perception of stereoscopic imageries enabled by the binocular vision. In addition, the nature of S3D images is fascinating and compelling enough that stereoscopic 3D applications have also been developed for many commercial and entertainment systems.

2.2 3D Displays

The history of 3D display systems have been well studied in [Okoshi, 1976] and hence will not be repeated here. Detailed reviews on traditional stereoscopic displays (requiring viewing glasses) can be found in [Valyus, 1966, Okoshi, 1976, Lipton, 1982, Lane, 1982, McAllister, 1993]. Reviews on new 3D technologies such as auto-stereoscopic (no glasses) displays, head-tracked displays (head tracker mounted on the display) and head-mounted displays (viewer mounting the display on his/her head) are available in [Eichenlaub, 1998, Morishima et al., 1998, Sanyo, 1997, Schwerdtner and Heidrich, 1998, Trayner and Orr, 1996, Woodgate et al., 1998, Urey et al., 2011]. The benefits such as more viewing freedom offered by those new technologies will be discussed later in this section.

Figure 2.8 presents a classification for current 3D display technologies as well as an outline for this section. While there are many different types of 3D displays, they can be generally classified into five main categories: stereoscopic displays, auto-stereoscopic displays, volumetric displays, holographic displays, and head-mounted displays. Most of the stereoscopic, auto-stereoscopic, and head-mounted displays have the problem of exhibiting the so-called “Vergence-Accommodation (VA) Conflict”, which is considered a great contributor to S3D visual discomfort (explained in detail in Section 3.2.1). Volumetric and holographic displays eliminate the VA conflict by reproducing correct accommodation/focus cues. More detailed information about how each technology works is provided in the following subsections.

2.2.1 Stereoscopic Displays

Stereoscopic displays, as illustrated in Figure 2.9, require viewers to wear a special viewing device that makes sure the left and right images are seen by the correct eye. LCD shutter glasses are popular for CRT displays [AG, 1999, Agrawala et al., 1997, Lipton, 1997] and polarising glasses are normally the choice for projection displays. As this type of 3D display technology has been well analysed in literature, only a brief summary of its three major types: Colour-Multiplexed, Polarisation-Multiplexed, and Time-Multiplexed, is presented below.

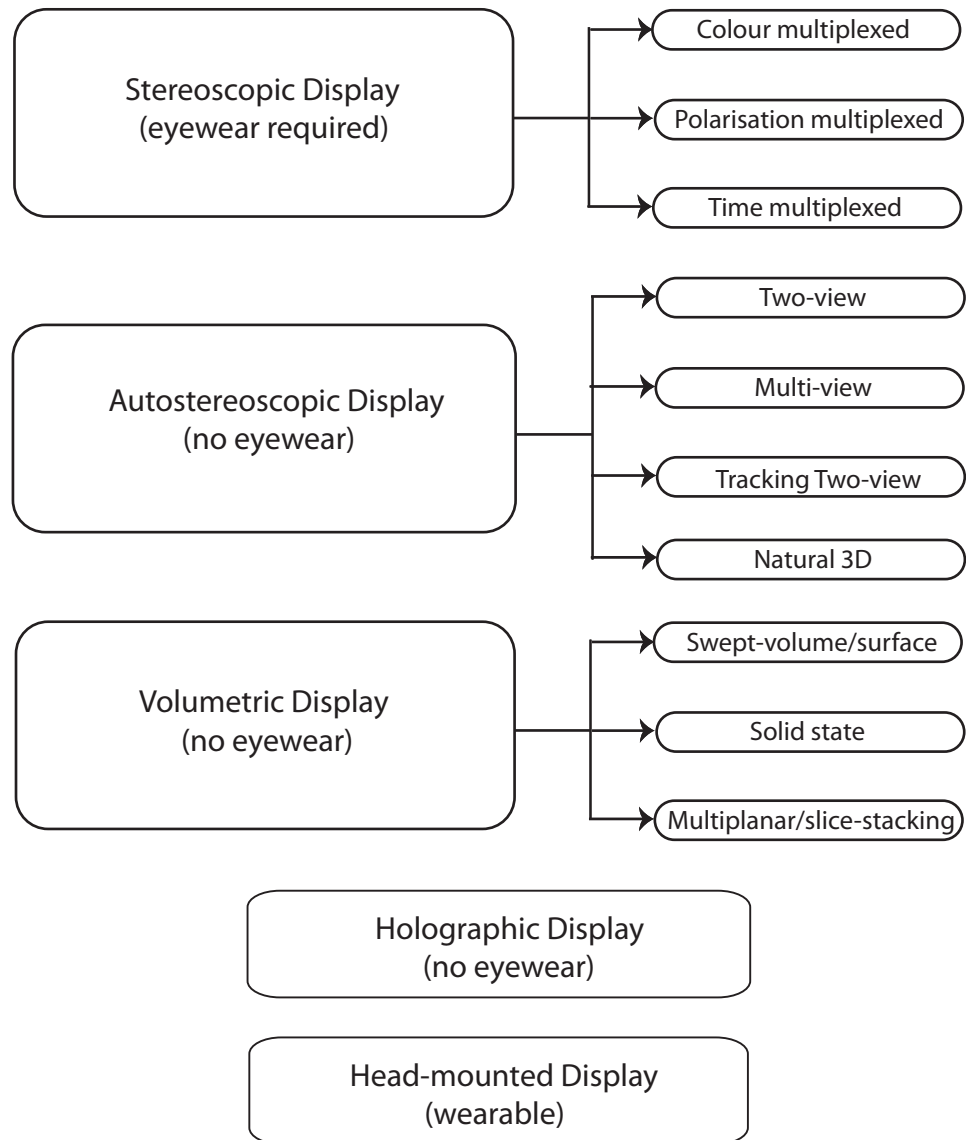


Figure 2.8: Organisation of the Section

Colour-Multiplexed Approach

The Colour-Multiplexed approach merges left and right images using a complementary colour coding technique. According to Urey *et al*'s survey [Urey et al., 2011], the understanding of this method dates back to the mid 1800s. Anaglyph glasses approach is the most typical application of this method. Viewers wear a pair of anaglyph glasses, which are colour-coded similarly as the images so that the left and right images are only available to the corresponding eyes. Red colour is the most common choice for the left image and eye whilst Cyan is used for the right image and eye, as demonstrated in Figure 2.10. Both the advantages and disadvan-



Figure 2.9: Stereoscopic Displays [Holliman, 2005]

tages of this method are obvious: the colour-multiplexed approach is inexpensive and convenient since it can be achieved with the most basic colour video equipment; however, this method does suffer seriously from the colour information loss and high degree of ghosting and crosstalk [Woods and Rourke, 2004, Ideses and Yaroslavsky, 2004, Ideses and Yaroslavsky, 2005]. Crosstalk is an artifact where one eye sees a portion of the image intended for the other eye. Ghosting is the visible crosstalk. Crosstalk and ghosting are deemed as one of the most serious problems associated with stereoscopic imaging [McAllister, 1993, Woods and Tan, 2002, Ronfard and Taubin, 2010]. We will explain them in detail in Section 3.2.3.

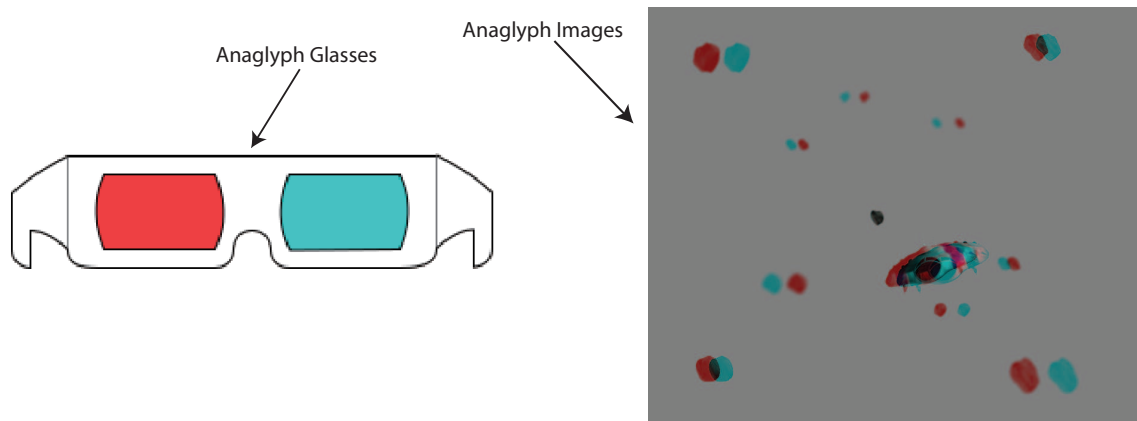


Figure 2.10: Anaglyph Method

Recently developed colour-multiplexed systems are: the ColorCodeViewer 3-D glasses and the Infitec system. The ColorCodeViewer 3-D glasses, designed by

ColorCode 3-D (Lyngby, Denmark), is a pair of anaglyph glasses which are colour coded by blue and amber, and have the capability of displaying full-colour 3D images on the standard 3D display hardware [Sorensen et al., 2004] (see [Col, 2011] for more information). Although this technique has been applied in S3D games, movies, and mobile content production, the generated 3D images still have serious ghosting and crosstalk problems.

The Infitec system, developed by Infitec GmbH (Ulm, Germany), is a unique colour-multiplexed based system that can produce full-colour 3D images with great image quality and low crosstalk ($< 1\%$ of the entire spectral range) [Gadia et al., 2008, Jorke et al., 2009]. This system employs two full colour projectors to project left and right images onto a diffuser screen through a set of special filters called the Infitec filters. The Infitec filter is an interference filter which has its own narrow transmission bands corresponding to Red, Green, and Blue channels. The eye wear is equipped with the corresponding Infitec filters to separate left and right image. They further developed a wheel-based single projector system, eliminating the image alignment problem as well as the complexity and cost of using two projectors. The wheel-based single projector Infitec system projects alternating left and right images onto a matte screen. The projector is fitted with a rotating Infitec filter to distinguish the left and right images coming out of the projector.

Polarisation-Multiplexed Approach

In the polarisation-multiplexed approach, the left and right images are polarised orthogonally to each other. Both circular and linear polarisation methods can be used while the former allows more head movement before the noticeable appearance of ghosting and crosstalk [Urey et al., 2011]. Viewers need to wear eyewears that are fitted with corresponding polarisers to block the unintended image for each eye. This method also employs a special screen to maintain the State of Polarisation (SOP) of each stereo image pair. A fresnel based lenticular screen is normally used by rear projection polarisation-multiplexed systems to preserve the SOP whilst the silver screen is the common choice for front projection systems.

DaLite (Warsaw, USA) has released commercially available 3D screens using

the polarisation-multiplexed method. The 3-D Virtual Black screen is available for rear projection systems and the 3-D Virtual Grey is developed for front projection systems [DaL, 2011]. Both the screens are reported to be able to preserve 99% of the polarised light. The ZScreen developed by RealD (Beverly Hills, USA) is a breakthrough idea for S3D projection. This system uses a single Digital Light Processing (DLP) projector with a ZScreen mounted in front of it. The ZScreen has a linear polarising filter and uses two pi-cells to control the direction of the polarisation [Lipton, 2001, Lipton, 2007] (explained in detail in Section 3.1.2). The problem of the image alignment, increased power consumption and cost in two-projector systems is eliminated by the employment of a single DLP projector [Bogaert et al., 2008, Bogaert et al., 2009]. Other available single projector polarisation-multiplexed systems are: HDI 3D projectors by HDI-US Inc. (Los Gatos, USA) [HDI, 2011] and the CF3D by LG (Seoul, South Korea) [CF3, 2011].

The advantages of the polarisation multiplexed approach are: the generated 3D images are of high quality and resolution; the passive polariser glasses are inexpensive and easy to use. The drawbacks of this method include: possible appearance of visible ghosting and crosstalk due to imperfect SOP preserving screens and polarisers [Brubaker, 2009]; a small percentage of the luminance of 3D images is lost due to the light polarisation.

Time-Multiplexed Approach

Time-Multiplexed displays rely on the human persistence of vision to fuse the alternating left eye view and right eye view of a stereo pair into a single stereoscopic image. In this method, the left and right images are projected onto the screen in a time-sequential fashion at a high frame rate, typically twice of the screen refresh rate: $2 * 60 = 120\text{Hz}$. Viewers are required to wear active shutter glasses which are in synchronisation with the left and right eye images being projected. The liquid crystal shutter glasses block the view of the right eye when the left eye image is displayed and *vice-versa*. The synchronisation between the stereo pair and the shutter glasses can be achieved by either an infrared emitter (explained in detail in Section 3.1.2), or the DLP Link (developed by Texas Instruments, Dallas, USA). The DLP

Link encodes flashes of white light in between left and right eye images and the shutter glasses can then detect those flashes and decide when to switch from one view to another.

In 1995, an interactive time-multiplexed system with head-tracking ability, named The Responsive Workbench, was developed by the German National Research [Kruger et al., 1995]. More recently, both Mitsubishi (Tokyo, Japan) and Samsung (Seoul, South Korea) released rear projection time-multiplexed 3D TVs on the commercial market [Chinnock, 2009]. Lightspeed Design, Inc. (Bellevue, USA) developed DepthQ 3D projectors available in both polarisation and time multiplexed methods. The DepthQ HDs3D-1 3D projector is a time-multiplexed single 3D projector system [Dep, 2011a]. This projector uses a single DLP projector to produce High Definition (HD) 3D images viewed with liquid crystal shutter glasses. The Cave Automatic Virtual Environment (CAVE) also employed the time-multiplexed technique [cav, 2011, Cruz-Neira et al., 1992, Leigh et al., 2007]. The CAVE is an immersive Virtual Reality system where 3D images are rear projected onto the walls of a cube of 10ft * 10ft * 10ft. Viewers need to wear active shutter glasses to experience the immersive feeling offered by this system.

Time-multiplexed 3D systems can produce 3D images with high resolution, quality and luminance. However, they suffer major drawbacks of the more expensive and complex use of battery-powered active shutter glasses and increased video bandwidth required.

2.2.2 Auto-stereoscopic Displays

Auto-stereoscopic displays, illustrated in Figure 2.11, overcome the drawback that viewers have to wear eyewears to separate left and right pairs. They send the left and right images directly to the intended eye. The key advantage that auto-stereoscopic displays have over stereoscopic displays is the ability of providing greater viewing freedom: viewers can see the 3D image without the help from visual aids as long as their pupils stay inside the so-called “viewing windows”. Auto-stereoscopic displays create virtual viewing windows that are parallel to the display screens in physical space. The virtual viewing windows form the so-called “nominal viewing plane”

where viewers have the most lateral viewing freedom. These viewing windows are 3D shaped diamonds tapering away from the display screen as shown in Figure 2.12.



Figure 2.11: Auto-stereoscopic Displays [Holliman, 2005]

While there are many different auto-stereoscopic display technologies enabling this effect, they can be generally classified into the following four major categories: Two-view, Multi-View, Tracking Two-view, and Natural auto-stereoscopic displays.

Two-view Auto-stereoscopic Displays

As demonstrated in Figure 2.13, a two-view display produces two views (one for the left eye and one for the right eye) that are visible in two (left and right) viewing windows. This display uses different sets of pixels on the screen to display left and right images and each eye can only see the intended image through the corresponding viewing window. The viewer's eyes are required to be at the correct position within the ideal distance to the display screen in order to perceive 3D images. Viewers can move approximately 20 to 30 mm around the central viewing position before the 3D effect is lost.

There are mainly four different two-view auto-stereoscopic design approaches:

1. Parallax barrier method. This approach is probably the most straightforward way to implement a two-view display. The parallax barrier, composed of vertical columns of apertures separated by black strip masks, blocks the alternately interlaced left and right image columns, allowing light of image pixels only pass to left and right viewing windows. Examples of two-view displays

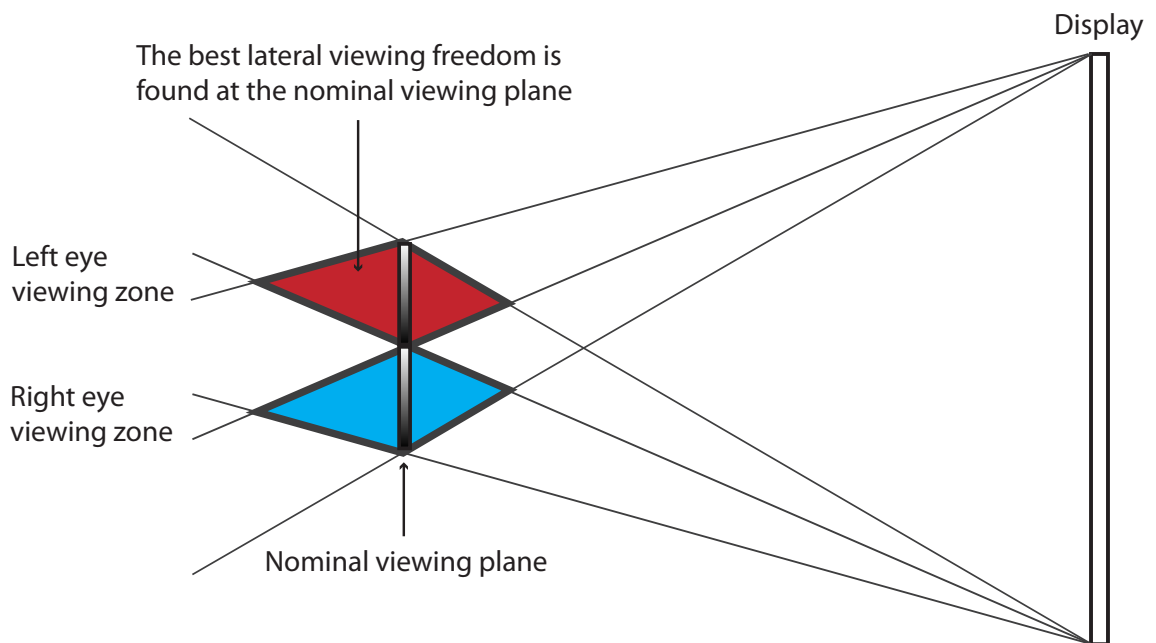


Figure 2.12: Viewing freedom in an auto-stereoscopic display [Woodgate et al., 1997]. The viewer can perceived a 3D image as long as his/her pupils stay inside the diamond-shaped regions

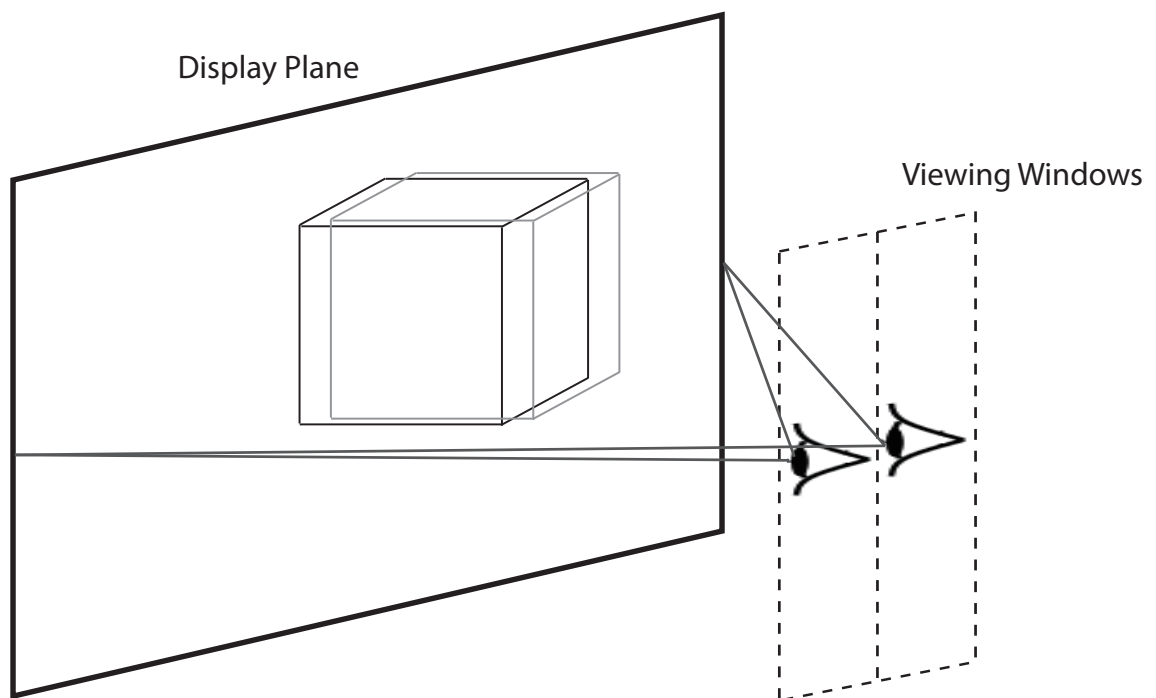


Figure 2.13: Two-view Auto-stereoscopic Display

are: [Woodgate et al., 2000, Yamamoto et al., 2002, Nishimura et al., 2007, Chen et al., 2009].

2. Lenticular element design. This method combines 2D displays with cylindrical lenses, which direct the diffused light from a pixel into a limited viewing angle separating left and right image pixels into left and right viewing windows respectively. [de Boer et al., 2007] is a good example of the lenticular system. More information about this approach is available in [Tsai et al., 2000, Johnson and Jacobsen, 2005, Lee and Ra, 2006].
3. Micro-polariser approach. This type of display uses micro-polarisers in combination with 2D displays to direct the light of left and right image pixels into their corresponding viewing windows. Examples are available in [Faris, 1994, Harrold et al., 2004].
4. Holographic components. This method employs illuminated Holographic Optical Elements (HOP), which are arranged in alternating horizontal strips, to create viewing windows [Trayner and Orr, 1996, Trayner and Orr, 1997]. More recent examples of this display use optical modules [Balogh, 2006, Balogh et al., 2007] or Spatial Light Modulators (SLM) [Stanley et al., 2004] in conjunction with holographic screens.

Two-view auto-stereoscopic displays provide benefits of generating 3D images with high resolution per view and low cost. Some of those displays are capable of switching between 2D and 3D displaying modes so that they can work as standard 2D displays when the 3D effect is not desired.

Multi-View Auto-stereoscopic Displays

The multi-view auto-stereoscopic display projects more than two (for example eight or nine) views at the same time into multiple viewing windows, as illustrated in Figure 2.14. Any two of these simultaneous views are arranged in the way that each pair of the left and right views forms a valid 3D image. Multi-view displays provide wide lateral viewing freedom, which allows viewers to move side to side around the

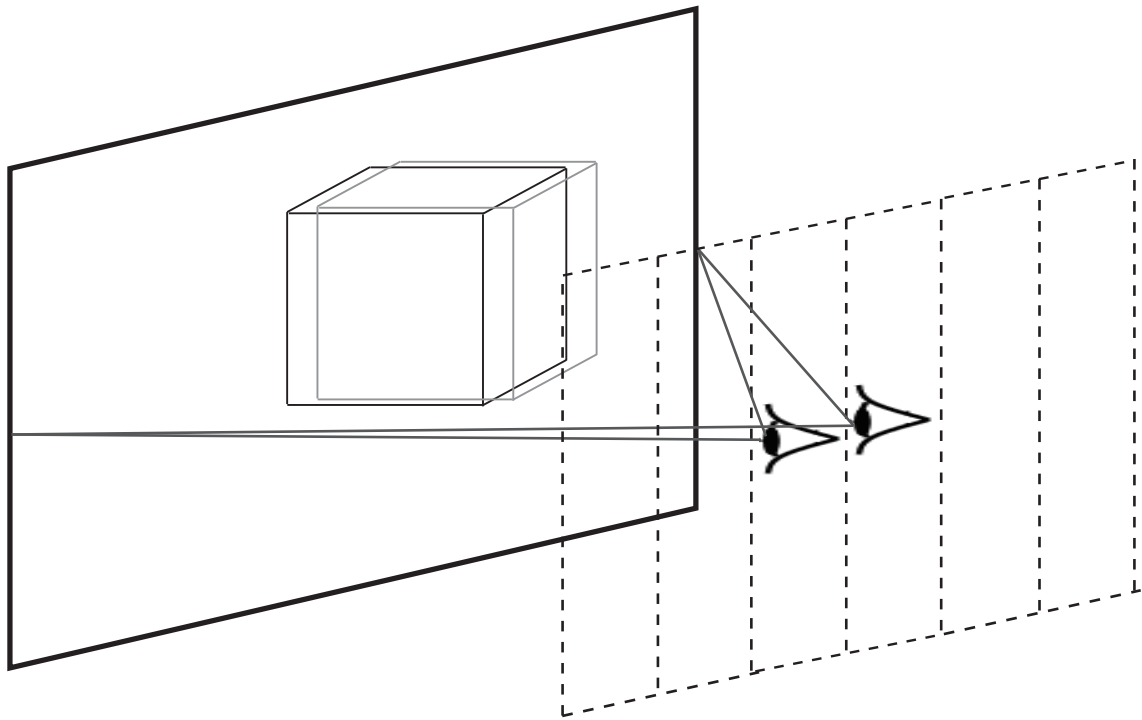


Figure 2.14: Multi-view displays create multiple viewing windows

display screen and can also support simultaneous viewing of multiple observers. The major drawback of these displays is the low resolution per each view as the total resolution of the display is divided by the number of views. These displays rarely have the function of switching between 3D and 2D modes.

Multi-view displays can be broadly categorised into the following types: Parallax Barrier [Choi et al., 2003, Mashitani et al., 2004, Ando et al., 2005, Sakamoto and Morii, 2006, Peterka et al., 2008, Lanman et al., 2010], Lenticular Array [van Berkel and Clarke, 1996, van Berkel, 1999], Time-Multiplexed [Moore et al., 1996], and Multi-Projector [Dodgson et al., 2000, Matusik and Pfister, 2004]. Different multi-view displays have different optimal viewing distances and viewing zones outside which stereoscopic images can not be perceived. Dodgson [Dodgson, 2002] provided a summary of equations for analysing viewing zones of different multi-view displays.

As discussed above, multi-view displays provide a wide lateral viewing range with the tradeoff of low image resolution as the total screen resolution is equally shared by each view. An interesting topic for multi-view display designers is what is the optimal number of viewing zones a multi-view display should produce so that

an ideal balance between viewing freedom and image quality can be reached? While Speranza *et al* claims that a better perception of smoothness can be achieved by large number of views [Speranza et al., 2005], the empirical analysis conducted by Hassaine *et al* suggests that a small number of view zones, each of which has a width of $\geq 1\text{cm}$, is required for good task performance in stereoscopic applications [Hassaine et al., 2010].

Tracking Two-view Auto-stereoscopic Displays

A tracking two-view display, as shown in Figure 2.15, is a two-view auto-stereoscopic display fitted with a head tracking mechanism. The head-tracker can track the position of a viewer's head or eyes and direct the left and right views to follow the respective eyes, enabling the viewing mobility. This type of display not only inherits the two-view display's ability of generating high resolution images, but also allows great lateral, even perpendicular to the screen in certain systems, viewing freedom. In addition, more advanced systems with fast and accurate head trackers are able to simulate the "look-around" effect where viewers can observe the object from all directions.

Tracking two-view displays can be implemented using micro optics [Ezra et al., 1995, Woodgate et al., 1997], parallax barrier [Perlin et al., 2000, Sandin et al., Perlin et al.,], and even holographic elements [Trayner and Orr, 1996, Trayner and Orr, 1997, Haussler et al., 2008]. A special type of the tracking auto-stereoscopic display that supports multiple mobile observers simultaneously, such as the MUTED display and HELIUM3D display, has been developed. The MUTED system comprises a direct-view LCD and steering optics directed by a laser illumination source. The HELIUM3D generate images by employing a fast light valve to control the output of an RGB laser. Both systems have individual eye trackers and require a very high frame rate so that the generated 3D images can be delivered to several mobile viewers simultaneously without any noticeable flickers [Surman et al., 2008, Brar et al., 2010].

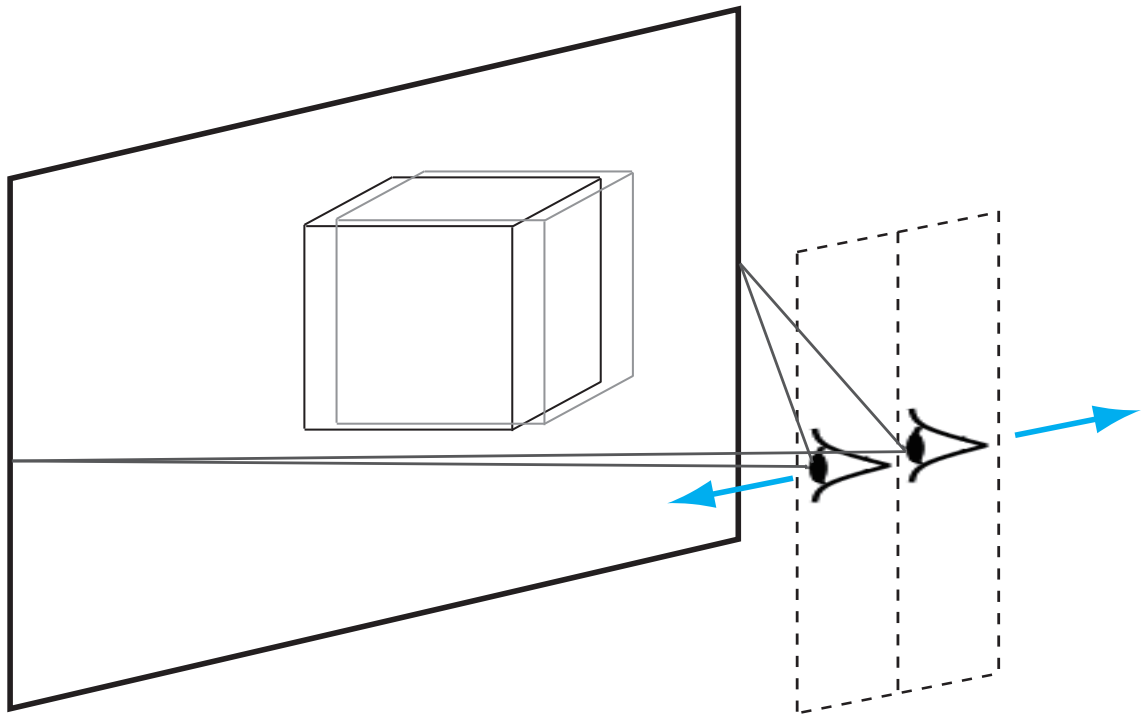


Figure 2.15: Tracking Two-view Displays

Natural 3D Displays

Natural 3D Displays [Takaki, 2009a, Takaki, 2009b] are auto-stereoscopic displays that can produce correct accommodation/focus cues and smooth motion parallax. There are two major types of natural 3D displays: Super Multi-View (SMV) displays and High-Density Directional (HDD) displays. A SMV display is a multi-view auto-stereoscopic display where the horizontal pitch between viewpoints is smaller than the diameter of the pupil, allowing more than one ray coming from the same point in space to enter the pupil simultaneously. A HDD display generates high-density directional rays whose angle pitch is so small ($0.2^\circ - 0.4^\circ$) that more than one ray can pass through the pupil at the same time, as illustrated in Figure 2.16.

Both the SMV and HDD systems produce solid visible image pixels when multiple rays intersect/converge in front of or behind the display so that viewers are able to physically accommodate or focus on those pixels, eliminating the problem of the Vergence/Accommodation Conflict. This is the reason that natural 3D displays are considered to have the potential to effectively function as volumetric displays. Examples of SMV displays are discussed in [Takaki, 2001, Takaki, 2009b]. [Takaki,

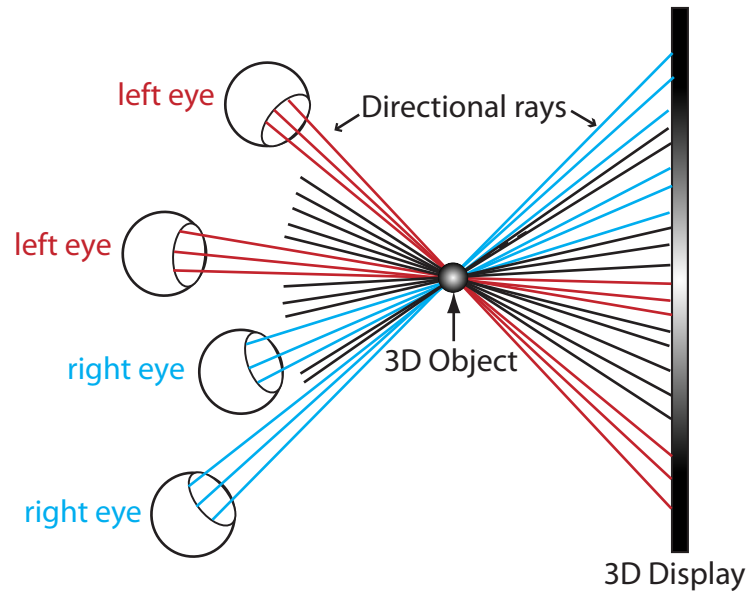


Figure 2.16: A 3D object displayed with high-density directional rays by a high-density directional display

2005, Takaki, 2006, Kikuta and Takaki, 2007] provide information about prototypes of HDD displays.

Natural 3D displays are in prototype only since they must employ multiple projectors or optical modules to generate 3D images with satisfactory quality, which requires a very complex design of display panels and extremely large data bandwidth. Research on how to make commercially-available natural 3D systems and the optical number of discrete light beams a natural 3D display should produce is needed.

2.2.3 Volumetric Displays

Volumetric displays are unique in that they do not simulate depth on a planar display screen through a variety of visual effects as other types of 3D displays do, but instead volumetric displays use volume pixels or voxels to actually reproduce depth that occupies a well-defined 3D volume by emitting, scattering, or relaying light in physical space [Holliman et al., 2011]. Volumetric displays are considered to be a promising solution to S3D viewing problems because of their ability to provide all the correct depth cues including accommodation without the help from visual

aides. There are mainly three different categories of volumetric displays: Swept-volume or Swept-surface volumetric displays, Solid-state volumetric displays, and Multiplanar or Slice-stacking volumetric displays.

Swept-volume/surface Displays

Swept-volume/surface volumetric displays reproduce 3D images by projecting 2D images (slices) onto a rotating or reciprocating surface designed to sweep the entire volume and rely on the human persistence of vision to fuse a time-series of volumes into a single 3D image. The Perspecta Display [Favalora et al., 2001, Favalora et al., 2002, Chun et al., 2005] designed by the Actuality Systems is a typical example of the swept-volume display. Perspecta creates spatial 3D images of approximately 100 million voxels in a transparent 10''(25 cm)-diameter dome. It employs three Digital Light Processing (DLP) engines to project a sequence of 2D patterns onto a omnidirectional diffuser screen rotating at 900 rpm or above. More detailed information about the Perspecta display can be found at [Per, 2011]. A couple of swept-volume displays with the capability of reproducing occlusion, complex reflections, and other viewer-position-dependent lighting effects, which are often mistakenly deemed impossible for volumetric displays, have been developed. In Cossairt *et al*'s Perspecta Display [Cossairt et al., 2007], they used mylar to replace the rotating diffuser screen. A tented brushed-metal surface was employed by the display designed by Jones *et al* with better rendering softwares [Jones et al., 2007].

Solid State Displays

Solid-state volumetric displays reconstruct 3D images without any moving devices, but instead they fill the display volume with a medium which can be excited to become emissive, so that voxels are invisible and transparent in the *off* state but opaque and luminous when switched on. This so-called "*Solid State*" process such as the two-step upconversion is described fully in [Lewis et al., 1971, Downing et al., 1996]. According to Chekhovskiy and Toshiyoshi's survey [Chekhovskiy and Toshiyoshi, 2008], a number of different substrates can be used for the solid-state volumetric displays such as water and air [Kimura et al., 2006, Saito et al., 2008]. One

example of this solid-state volumetric display is introduced in [Lewis et al., 1971], which has two perpendicular lasers to excite certain areas in a gas-filled volume.

Multiplanar/Slice-stacking Displays

Multiplanar/Slice-stacking volumetric displays do not use any moving devices or emissive medium, they project a sequence of 3D image slices onto multiple (a stack of) optical elements positioned at different depths. Louis Lumiere provided the description of the first known example of slice-stacking displays, which created a photosculpture by stacking photographs of successive optical tomographies [Hassaine, 2010]. A typical example of this display is the DepthCubeTM Display system developed by the LightSpace Technologies [Sullivan, 2004]. DepthCube also uses a high-speed DLP projector to project more than 15 millions physical voxels onto a Multiplanar Optical Element (MOE), which is composed of a stack of 20 liquid crystal scattering shutters acting as the 3D volume for projection (for more information see [Dep, 2011b]).

Advantages and Disadvantages of Volumetric Displays

The advantages of volumetric displays are obvious, such as the ability of providing correct accommodation cues, a wide range of available viewing positions allowing viewers to walk around the display, etc.. A major drawback associated with volumetric displays is the data bandwidth problem since volumetric displays require a large number of views rendered at the same time, and consequently are rather expensive. An open question for volumetric displays is at what point a natural 3D display, which is potentially capable of providing correct accommodation cues, will become practically equivalent to volumetric displays.

2.2.4 Holographic Displays

Holography is a technique that records a light field and reconstructs it later when the original light field is no longer present. During the process of holographic exposure, a hologram forms microscopic interference fringes to record information about the light wavefronts scattered off an object. When the developed holographic film is

illuminated appropriately (same light source and same angle as it is recorded), the microscopic interference fringes act as a complicated diffractive lens, reconstructing all the characteristics of the original light source (see [Hariharan, 1996] for more detailed description, definitions of interference, diffraction, etc.). Holograms are able to reconstruct the original light source so exactly that humans can have difficulties in distinguishing them from the scenes in the natural world [Halle, 1994]. Holographic displays also share volumetric displays' ability of providing all the correct depth cues including accommodation. The term "Hologram" or "Holographic Display" has often been erroneously used to describe displays that are merely vaguely three-dimensional. To be clear, Holographic displays only refer to optical devices that are image-bearing diffractive [Halle, 1997]. Other 3D displays ought not to be called holographic displays.

Limitations of Holographic Displays

There are two major difficulties in holography: (1) During the process of recording, even the slightest movement of the recording devices or in the scene can seriously damage the quality of the hologram due to the nature of this technique. (2) In order to provide correct accommodation/focus cues and smooth motion parallax, holographic displays employ a Spatial Light Modulator (SLM) to provide the required ultra high resolution to match the wavelength of light. The spatial frequency of a typical hologram is more than 1500 line pairs per millimetre [Halle, 1997] ($\sim 1 \mu\text{m}$ of pixel pitch according to [Takaki, 2009a]). Also the number of pixels rendered by the SLM must be increased proportionally with the increase of the screen size of a holographic display, which makes the generation of a hologram larger than a few centimetres very difficult. The nature of diffraction and light also imposes problems in realising holographic displays, such as the huge calculation requirement to calculate holograms at an interactive rate [Balogh et al., 2005], modulation and speckle noise, and narrow fields of view. These are the reasons that holographic displays still remain to be a challenging research field.

Like volumetric displays, holographic displays also suffer from data bandwidth problems. The Horizontal Parallax Only (HPO) holography has been invented to

reduce the number of pixels by eliminating vertical parallax, which does not affect the sense of depth, and generating only horizontal parallax. A few good ideas in implementing the HPO holographic display have been proposed, including Takaki and Hayashi's resolution redistribution HPO system [Takaki and Hayashi, 2008] and the horizontal scanning holography technique introduced by Takaki and Okada [Takaki and Okada, 2009]. Holographic stereogram is another technique designed to further ameliorate the data bandwidth problem. A holographic stereogram uses a relatively small number of discrete perspective views to approximate the continuous parallax of a hologram. The number of the employed discrete perspective views has to be large enough so that a continuous range of perspectives can be perceived. However, holographic stereograms have their own drawbacks, such as inter-view aliasing artifacts caused by inadequate or improper wavefronts sampling and incorrect accommodation [Halle, 1994].

2.2.5 Head-mounted Displays

Head-mounted or Helmet-mounted stereoscopic displays (HMD) [Pastoor and Wopking, 1997], illustrated in Figure 2.17, differ from other types of 3D displays in that viewers are required to wear them on the head to perceive the displayed 3D images. An immediate advantage of this approach is the viewing mobility without losing the perception of 3D images. A typical HMD 3D display employs two small screens (one for each eye) with relay lenses to direct the left and right images into the corresponding eyes separately. As HMD displays have large viewing angles, typically up to 80° (vertical) and 120° (horizontal), and block the sight of the natural environment, they can provide viewers with a true feeling of immersion. The first commercial head-mounted 3D display, HMZ-T1, was released on the consumer market by SONY at the time this thesis was written (see [HMZ, 2011] for more information).

The first attempt to create an immersive display was the Sensorama Simulator developed by Heilig *et al* in the late 1950s [Heilig, 1962]. Nowadays, HMD displays are widely exploited in military, engineering, medical, and commercial fields. In particular, Augmented Reality (AR) systems, which augment viewer's perception of the real world by using the so-called "see-through" approach: superimposing



Figure 2.17: Head-mounted Display [HMD, 2011]

Computer-generated Imageries (CGI) on real scene images captured by video cameras, can substantially benefit from HMD 3D displays. A few good see-through HMD displays have been proposed for AR systems: Ferrari *et al* designed an HMD 3D system which can superimpose computer-generated tomographies on live images of real patients to help surgeons study the preoperative radiological exams [Ferrari et al., 2009]. The HMD display developed by Thompson can present virtual 3D tunnels for paratroopers to navigate through the sky [Thompson, 2005].

A challenging research topic for designing HMD displays is to achieve high resolution and large Field of View (FOV) at the same time. Optical tiling, which refers to generating small high resolution inset images with low resolution background images, has been proposed to satisfy the requirement of simultaneous high resolution and large FOV [Howlett, 1992, Rolland et al., 1998]. Sensics Inc (Columbia, USA) provides HMD 3D displays with HD-quality resolution and panoramic (up to 180° diagonal FOV) viewing. The core element of those displays is the piSight unit which employs a concave display array of 12 micro displays to form a wrap-around image for each eye [Sen, 2011].

As HMD systems display left and right images separately, any misalignment or mismatch issues between the left and right image channels can introduce serious visual discomfort for viewers [Cakmakci and Rolland, 2006]. In addition, like most of the stereoscopic and auto-stereoscopic displays, HMD 3D displays also suffer from the Vergence/Accommodation Conflict. In order to solve the VA Conflict problem,

HMD 3D systems with adjustable focal planes have been developed [Takahashi and Hirooka, 2008, Liu and Hua, 2009].

2.3 Summary

In this chapter, we presented relevant background information in two topics: the human visual system and the state of the art in 3D display technologies.

In the section of the human visual system, the principle of how the human visual system perceives binocular depth in the real world and in electronic stereoscopic images was discussed. Other depth cues including monocular cues and motion parallax were also introduced. In addition the benefits that the binocular vision is able to provide over monocular vision in the real world were provided.

Regarding the development and applications in 3D display systems, we reviewed most of the available stereoscopic and auto-stereoscopic display designs, ranging from the traditional 3D displays which require viewers to wear 3D glasses to perceive the 3D images to the most state of the art natural auto-stereoscopic displays which can produce correct accommodation/focus cues without the help from visual aids. Technologies such as volumetric, holographic, and head-mounted stereoscopic displays were also briefly discussed.

To summarise, stereoscopic displays provide viewers with a more natural viewing experience as they can imitate the real-life viewing process by presenting a different view to each eye and displaying images in three-dimensional. However, due to the limitations in the human visual system and stereoscopic displays designs, most of the current 3D displays still exhibit the Vergence-Accommodation (VA) Conflict which greatly contributes to the S3D visual discomfort. In the remainder of this thesis, analysis of this VA conflict and how it affects the S3D viewing especially in stereoscopic cinematography will be provided alongside with our solutions to this problem.

Chapter 3

Stereoscopic Cinematography

“Stereoscopic Cinema, the last great innovation.” Lenny Lipton [Lipton, 2007]

Since the release of James Cameron’s *Avatar* in year 2009, there has been an explosion of producing and projecting movies in S3D format: more than 40 mainstream S3D movies were released in 2010 and about 60 S3D movies were scheduled to be released in 2011 [3dm, 2011]. Stereoscopic cinema has become the new great hope of the industry as recently released S3D films were producing three-times the revenue per screen of the simultaneously released 2D films [Digest, 2007, Mendiburu, 2009]. However, the resurrection of the S3D cinema has been coupled with adverse effects associated with S3D visual discomfort. In this chapter, we provide a historical review regarding to the cinema technology, explore the reasons for the recent renaissance of stereoscopic cinema, analyse technical issues in S3D, discuss research problems brought by applying *Cinematic Storytelling* techniques in stereoscopic cinematography, and finally, summarise well-known key factors in producing a comfortable viewing S3D movie and discuss remaining research challenges in stereoscopic cinematography.

3.1 Cinema Technologies

3.1.1 Review of Cinema Technology Evolution

The technical evolution in cinema, as illustrated in Table 3.1, is a huge topic and we could only scratch the surface here. In this section, we provide a brief discussion about the prior cinema technical innovations such as sound, colour, and widescreen along with the present introduction of the stereoscopic cinema.








Technology	Intro.	Practical Impact	Artistic Impact
 Motion	1895	Required new infrastructure to exhibit	Created new visual narrative techniques
 Sound	1927	Increased both production and exhibition costs	Dialog became important for storytelling
 Colour	1935	Only increased production costs	Created new techniques for emotional expression
 Widescreen	1952	Increased both production and exhibition costs	Improved movie viewing experience
 3-D	1952	Made both production and exhibition more difficult and expensive	Failed to significantly impact the nature of film narrative
 Digital Projection	2000	More reliable and simpler than film projection. Required initial investment.	No obvious difference to the viewers
 3D Digital Projection	2005	More reliable and simpler than 3D film projection. Required initial investment.	Improved viewing experience and extended palette for emotional expression

Table 3.1: Innovations in Cinema Technology [Lipton, 2007]

Sound

There are many similarities between the introduction of sound and the introduction of the stereoscopic cinema. Both were explored decades before commercial acceptance; both required the synchronization of two machines; both required initial investments from the producers and exhibitors; and both imposed a rethinking of filmmaking techniques such as production design, cinematography, and production and exhibition pipeline issues.



Figure 3.1: Vitaphone System [Vit, 2011]

Sound movies were eventually made commercially practical by the invention of *Vitaphone* system, shown in Figure 3.1. The Vitaphone system was introduced by Warner's Bros in 1927. It used a projector that had a mechanical interlock with an attached phonograph. The movie *The Jazz Singer*, including a two-minute dialog and several synchronised songs played by the vitaphone system, broke the box-office record and established Warner Bros as a major force in the Hollywood.

Colour

Although just about every movie we watch today is in full colour, and as they say, the colour movies are perfected. Colour technology took a long time to prevail and become one of the most important cinema storytelling techniques. Colour films did

not require any upgrade cost from the exhibitors as they could play them using the same projectors and screens that had been used for the black-and-white movies. However, colour movies were regarded with derision in their early days (1920s) of introduction mainly due to the bad quality of the colour images. The studios and distributors had to make a significant investment in exploring colour photography and prints in order to balance between capturing good quality full-colour images and making good quality prints relatively low-cost for distribution. Hence, colour movies often demanded an increase in tick prices in the early days of introduction, and eventually, it commanded a premium at the box office and became the dominant force in the industry. Interestingly, there is no difference today in the cost of producing a colour movie comparing with that of a black-and-white movie.

Widescreen

In the early 1950s, the widescreen technology such as *Cinerama* and *CinemaScope* was introduced. **Cinerama** was a widescreen process with stereoscopic sound system which simultaneously projected images from three synchronised 35 mm projectors onto a huge, deeply-curved 150° of arc screen, shown in Figure 3.2. Cinerama was presented to the public as a theatrical event, with reserved seatings and printed programs. **CinemaScope** was another widescreen movie format used in the 1950s. It did not require multiple projectors like the Cinerama did. Instead it added *Anamorphic* lenses in front of the projector to project film up to a 2.66:1 aspect ratio onto a big screen. This process is called anamorphosis, illustrated in Figure 3.3, where a curved lens placed in front of an ordinary camera squeezed a large field of vision into the frame. Another lens, affixed to the projector, would reverse the process, spreading out the image onto an extended screen. Although CinemaScope was shortly made obsolete by new technological developments, the anamorphic presentation of films initiated by CinemaScope has continued to this day. Nowadays almost all movies are shot in a wide aspect ratio. The conventional Edison's aspect ratio of 1.33:1 is the exception (analog TV and IMAX).

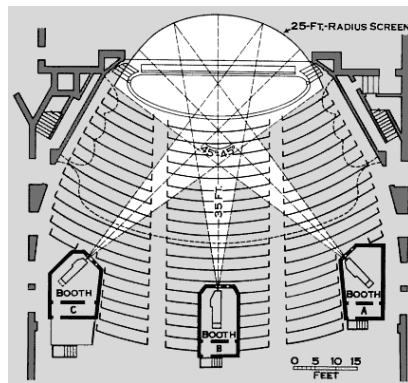


Figure 3.2: Cinerama Diagram

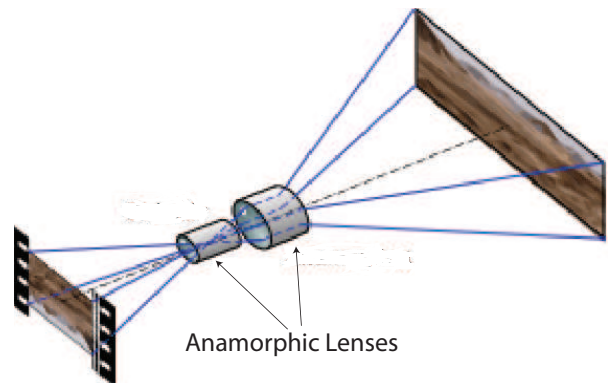


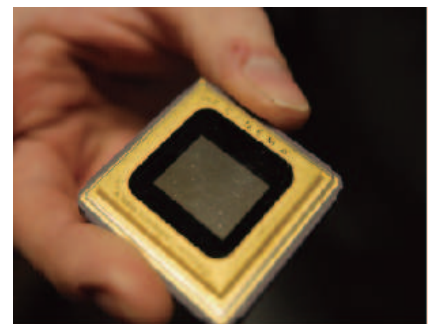
Figure 3.3: Anamorphic Lenses

Digital Projection

The cinema industry has been going through a fundamental change from the conventional film infrastructure to a digital system since the year 2000. The digital cinema systems defined by the Digital Cinema Initiatives (DCI) includes four major processes: Digital Capture, Digital Post-production, Digital Mastering, and Digital Projection. A movie can be shot at high definition by digital cameras, distributed via hard drives, optical disks or satellite and projected using a digital projector instead of a silver-based 35mm film projector.



DLP Projector



DMD Chip

Figure 3.4: DLP Projectors

Digital Projection is still in its early days of introduction. Approximately 4000 out of 135,000 theaters worldwide have deployed digital projectors [Lipton, 2007],

virtually all of which are using the Digital Light Processing (DLP) projector based on the digital micro-mirror device (DMD) developed by Texas Instruments (Figure 3.4). Given the obvious advantages the digital projection has over the film projection such as film transport and frame damage, the pace of the deployment of digital projection system has been accelerating for the past several years.

3.1.2 Stereoscopic Cinema

S3D Film Projection

The first ever commercially successful full-colour stereoscopic film was projected in the U.S. in 1940 [Lipton, 2001]. More than 60 stereoscopic films were shot in the early 1950s. However, many of those were not released in stereoscopic format because of the technical problems such as the synchronisation of the two projectors and scratched and damaged films. Stereoscopic movies projected by films and two projectors did not prevail for long and failed to significantly impact the film nature at the first attempt.

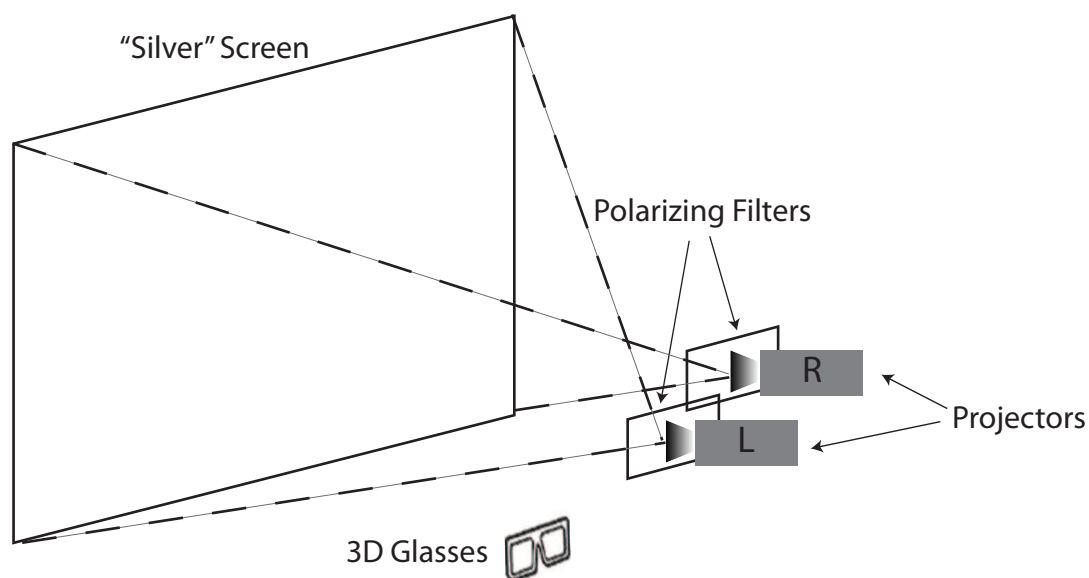


Figure 3.5: Stereoscopic Film Projection

The projection process used at that time is demonstrated in Figure 3.5. There were two projectors in the booth for changeover from reel to reel, providing an

opportunity to modify the setup to run interlocked left and right projectors. The polarized light method was used with filters over the projectors and audiences' eyes. A special screen, the so-called "Silver" screen was used to conserve the polarization. This is still the basic principle the IMAX 3D cinemas are using to project the stereoscopic films today.

S3D Digital Cinema

With the development of the digital projection technique, one could make the case that stereoscopic movie is the "ultimate" application of cinema technology. However, as mentioned above, stereoscopic projection based on the two-projector system could be very problematic. The invention of the single-digital-projector system, based on the *ZScreen*, eventually made the production of an enjoyable stereoscopic format movie possible.

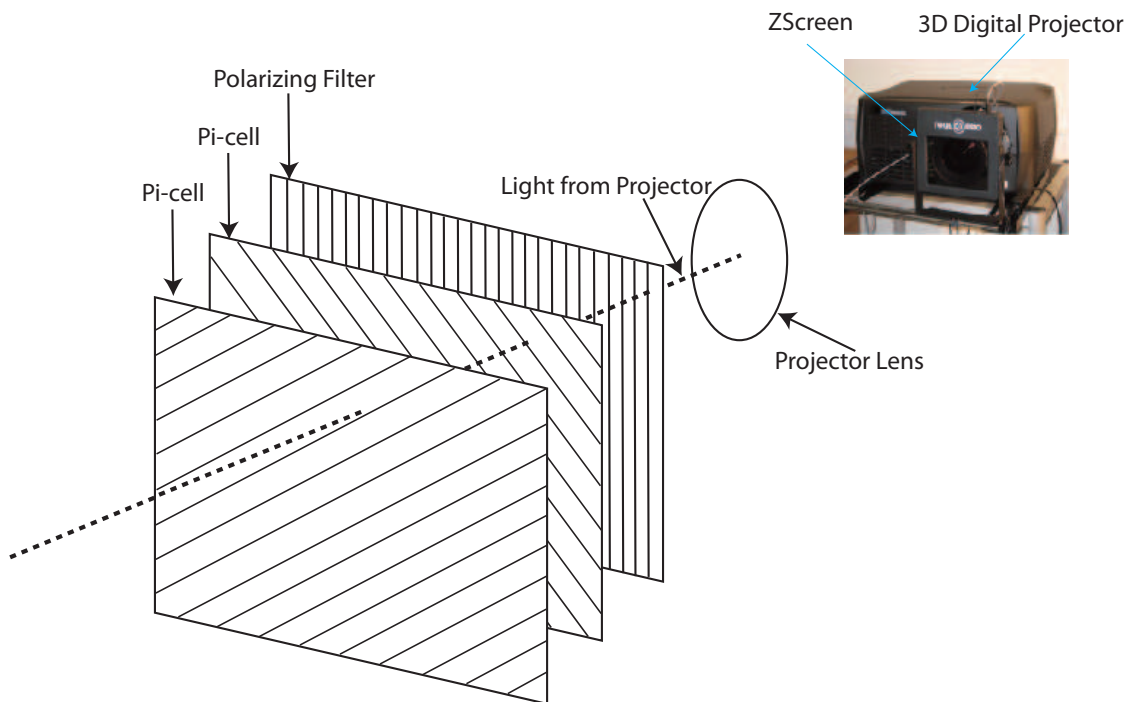


Figure 3.6: 3D Digital Projection System

Figure 3.6 illustrates the single 3D digital projector with ZScreen. For most of the modern 3D projection systems, only a single 3D projector with a ZScreen mounted in front of its lens is required. The ZScreen is capable of switching the characteristics

of polarized light at the same frame rate of the projector. The ZScreen is made of one linear polarising filter and two pi-cells in optical series. The pi-cells switch on and off to create left and right handed circularly polarised light synchronising with the left and right images. Since the frame rate of the projector is twice the frame rate required, viewers will be able to watch the movie without any graphical glitch.

Other Projection Methods

Using polarised light to separate the left and right images has always been the prevailing method for the projection of stereoscopic movies, both for the conventional two-projector film projection and the recently developed single-digital-projector system. In addition to this so-called “polarisation” method, three other techniques that had been explored for the theatrical 3D movie projection are worth mentioning: *Anaglyph*, *Vectograph*, and *Eclipse* technique.

Anaglyph

Anaglyph technique, as illustrated in Figure 2.10 on page 25, refers to the method of separating the left and right images using complementary coloured images. Viewers were required to wear the glasses using the similar complementary coloured filters. One year before the release of the first full-colour stereoscopic movie, a similar movie was projected using the Anaglyph projection in the U.S. Since then, the Anaglyph technique had a long history of going in/out of fashion during the early days of the stereoscopic movie introduction. Both its strengths and drawbacks are quite obvious: it only requires a single projector, but entails the projection of the monochrome images, which were considered to be the main reason that stopped its commercial acceptance.

Vectograph

Vectograph process was invented by the Polaroid Corporation. Unlike Anaglyph, this technique provides full-colour image projection and it works by imbibing polarising dyes onto the two reels of left and right films which are then stacked together into a single film [Land, 1942]. The single film contains both the left and right images with each being able to polarise light, and when viewed through polarising

3D glasses, the viewer sees a three-dimensional image. This technology was never used for commercial stereoscopic movie projection. The most common image reproduced by this process is probably the Titmus Fly Stereotest (Figure 3.7), which is used by Optometrists and ophthalmologists to determine if subjects have normal stereoscopic vision.



Figure 3.7: Titmus Fly Stereotest

Eclipse Technique

Eclipse technique was first proposed for slide projection in 1855 [Judge, 1950]. The principle of this method is to alternately block and pass the images for left and right eyes. In 1922, Laurens Hammond and William F. Cassidy invented the first commercially available eclipse projection system, *Televue* [Hammond, 1924], which was used to screen the movie *The Man From M.A.R.S.* in the only theater that deployed the equipment, the Selwyn Theatre in New York City. In the televue theatre as shown in Figure 3.8, there was a gooseneck mounted on the back of every seat. Mounted on top of the gooseneck was a spinning shutter electrically synchronised with the projector's shutter. At the time the viewer's left eye was blocked, so was the left projector's lens. As the viewing shutters continued to rotate, the left view and projector's lens was unblocked and so on and so forth. The repetition rate needed to be high enough to satisfy the condition of critical flicker frequency in order to achieve a flicker-free stereoscopic movie projection.



Figure 3.8: Teleview System [Tel, 2011]

Eyewear

Active Eyewear

Active eyewears are battery-powered, liquid crystal shutters that run synchronously with the field rate of the video feed. The synchronization between eyewears and video feed is achieved by an infrared emitter, which monitors the video signal and identifies the coded infrared pulses that distinguish the left and right images [Lipton, 1990]. The eyewear contains a detection diode of infrared which controls the shutters to occlude or transmit. An active eyewear product *CrystalEyes*, developed by StereoGraphics corporation, is commonly used in today's Virtual Reality applications such as the CAVE [cav, 2011]. Interestingly, Lenny Lipton and his team conducted an experiment of viewing high-parallax 3D images with *CrystalEyes* glasses. Their results suggested that the *CrystalEyes* glasses do not allow light leak and it was the CRT projector, working in conjunction with *CrystalEyes* glasses, contributed to the appearance of ghosting [Lipton, 2001].

Passive Eyewear

Passive eyewears, also known as polarised glasses, contain a pair of orthogonal polarising filters. As each filter passes only the similarly polarised light and blocks the orthogonally polarised light, each eye only sees its separately polarised image. Although most of the 3D cinemas project stereoscopic movies using linear polarising light, research has shown that the circular polarisation method, using left and right-handed circular polarised light to select images, allows a lot of head movement before



Figure 3.9: 3D Eyewears

losing the stereoscopic effect [Walworth, 1984]. Comparing to the active eyewears, the dynamic range of passive eyewears is lower. However, the passive eyewear has its virtues such as better image quality and lower cost. In addition, there are other venues where cardboard or plastic-framed eyewears should be used instead of the more expensive active eyewears. Figure 3.9 illustrates an example of passive and active eyewears.

Summary

After languishing for more than a century, the evolution of cinema technology has finally come to the stereoscopic format movies. The future of the film industry is upon stereoscopic movies. Over the past several years, technologies regarding to the projection of stereoscopic films have been significantly improved and quite a few stereoscopic movies with decent quality have already been released. However, despite decades of exploration to perfect it, the stereoscopic filmmaking is still in its infancy [Lipton, 2007]. In order to make stereoscopic cinema an accepted part of the neighbourhood theatrical experience, one must start with a basic understanding of the unique characteristics of the stereoscopic cinema and their implications, which will be addressed in the next section.

3.2 Technical Issues of the Stereoscopic Cinema

The “Resurrection of Stereoscopic Cinema” has been accompanied by concerns on adverse effects associated with viewing S3D content. A recent research survey reported that 30% of viewers experienced eye fatigue after watching an S3D movie [Ber, 2010]. A display manufacturer recently issued recommendation against watching S3D television as it could cause altered vision, headache, nausea, at extreme cases, convulsions if the viewer is in bad physical condition or pregnant [Shibata et al., 2011]. It is obvious that the ultimate prevailing of S3D technology is significantly hindered by those adverse effects regarding viewing discomfort. In this section, we review the causes and solutions of those adverse effects.

3.2.1 Vergence-Accommodation Conflict

The nature of the Vergence-Accommodation (VA) Conflict was discussed in Chapter 1. Here we focus on the viewing discomfort imposed by the VA Conflict.

It is a consensus in literature that VA Conflict *per se* causes the visual discomfort in viewing S3D content. However, the great majority of the literature supported that theory based only on the fact that viewers reported more viewing discomfort when viewing stereoscopic displays than when viewing non-stereoscopic 2D displays [Emoto et al., 2005, Häkkinen et al., 2006, Yano et al., 2002]. Unfortunately, this observation alone can not prove that the VA Conflict *per se* causes S3D viewing discomfort as there are several other substantial differences between viewing S3D displays and viewing non-stereoscopic displays, such as eyewears required to separate left and right views, the image misalignment [Kooi and Toet, 2004], distortions that take place in viewing S3D displays but not in 2D displays [Woods et al., 1993, Bereby-Meyer et al., 1999, Vishwanath et al., 2005], and the impact of head movements during viewing S3D images [Wartell et al., 1999]. To our knowledge, only the following two studies had allowed viewers to watch the same stereoscopic content with and without the conflict, providing convincing evidence that the VA Conflict in S3D indeed causes visual discomfort.

Sugihara *et al* proposed a unique stereoscopic display system called the 3DDAC

[Sugihara et al., 1999]. The 3DDAC can project virtual stereoscopic images at different focal distances through a movable relay lens. Therefore, the focus distance of the displayed 3D images can be easily controlled by moving the relay lens. This approach provided an excellent platform to test viewers' visual comfort with different VA relationships: (1) no conflict, the focus distance is equal to the vergence distance; (2) small conflict, the difference between the focus distance and vergence distance is small; and (3) large conflict, the difference between the focus distance and vergence distance is large. The results of their subjective assessment confirmed that the VA Conflict is indeed a factor that causes visual discomfort when viewing S3D images.

Akeley, Watt and Hoffman *et al* developed a novel volumetric display [Akeley, 2004, Watt et al., 2005b, Hoffman et al., 2008] and achieved correct or near-correct focus cues, eliminating the VA Conflict. This system was used to compare viewers' symptoms resulted by different VA relationships (no conflict, small conflict, large conflict, etc). Their subjective assessment results also proved that the VA Conflict itself does cause visual discomfort for viewers. However, a tradeoff of this method was that a significant portion of light that would be seen by viewers was absorbed by the system itself due to the use of multiple image planes and the depth-weighted image intensity filtering technique.

Numerous studies have been conducted on controlling the perceived depth to reduce viewing discomfort imposed by the VA Conflict. We discuss them in detail in Chapter 4.

3.2.2 Camera Configuration

The term "camera" here can either refer to the actual physical camera or a virtual camera created by computer graphics, representing the centre of projection which projects the scene into the virtual space. The virtual camera uses cartesian coordinates to describe a viewing frustum. The viewing frustum, as shown in Figure 3.10, is a rectangular pyramid frustum that defines the viewing volume in virtual space. The near and far clipping planes cut the frustum perpendicularly to the direction of viewing. A perspective projection matrix transformation is required to transform objects within the viewing frustum onto the display plane.

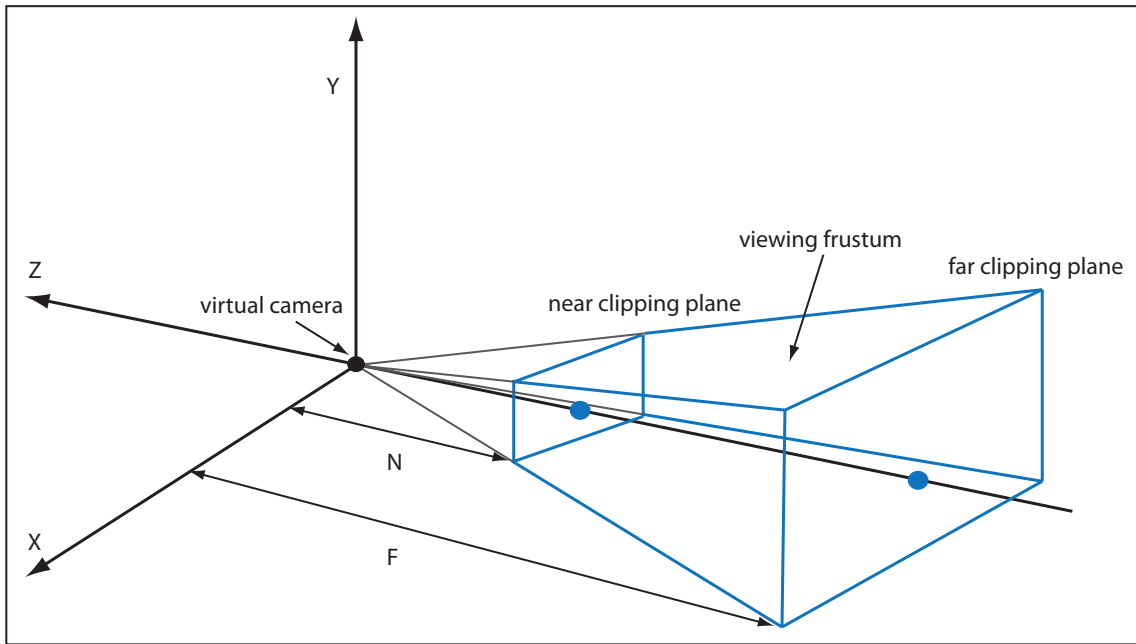


Figure 3.10: The canonical viewing frustum. Notations: N is the distance between the virtual camera and the near clipping plane. F is the distance from the far clipping plane to the virtual camera. Distances between the virtual camera to the left, right, top, bottom clipping planes also need to be specified. Only objects that are inside the viewing frustum are drawn on the display screen.

One of the most discussed issues regarding stereoscopic imaging is the camera arrangement. Figure 3.11 demonstrates the arrangements of two common approaches: the converged camera model and the parallel camera model.

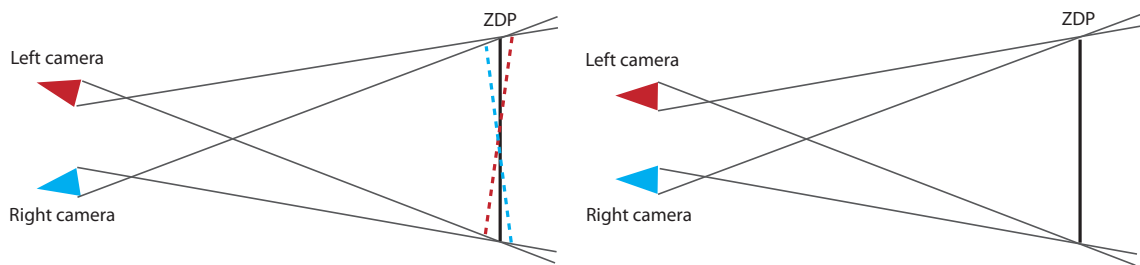


Figure 3.11: Converged camera model (left) and parallel camera model (right) [Holiman et al., 2006]

The converged camera, also known as the toed-in camera arrangement is shown in left part of Figure 3.11. People working in the real-world stereoscopic photography often prefer this approach as it is easy to setup the Zero Disparity Plane (ZDP)

by converging the left and right cameras. However, this rotation generates opposite keystone distortion [Woods et al., 1993, McAllister, 1993, Diner and Fender, 1993] in the left and right images. Keystone distortion leads to vertical disparity in the stereoscopic pair due to the images being located on different planes. Vertical Disparity, illustrated in Figure 3.12, is the vertical difference between two corresponding points in left and right views. In the left view, the left side of the image appears to be smaller than the image on the right view. In the right view, this effect is reversed.

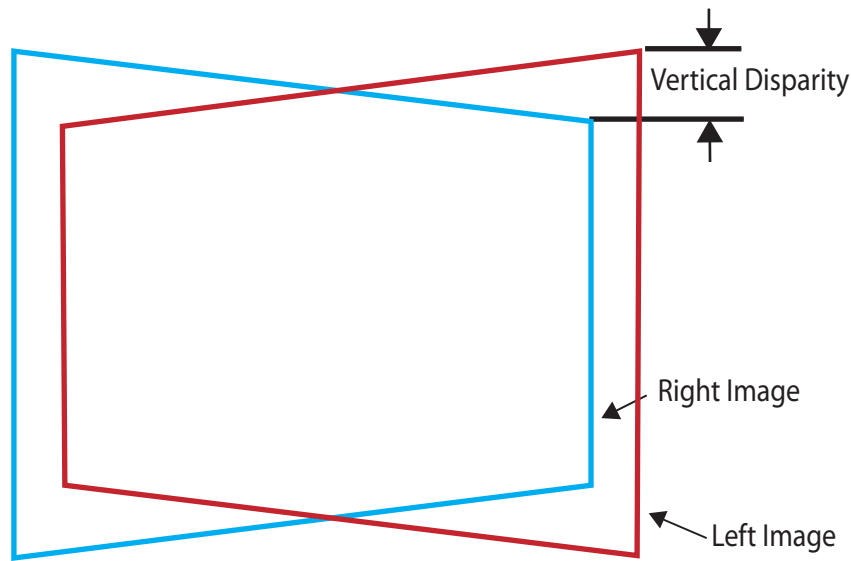


Figure 3.12: Vertical Disparity [Woods et al., 1993]

A parallel camera arrangement, illustrated in the right part in Figure 3.11, eliminates the problem of keystone distortion by parallelly aligning the axes of two cameras and placing the left and right images on coplanar image planes. Two approaches can be used to ensure the infinity is behind the display screen. The camera frustums can either be asymmetrical (shown in the right part in Figure 3.11) so that the cameras have coincident field width at the ZDP, or be symmetric but needs to be cropped to the region which is inside both left and right views as illustrated in Figure 3.13.

The parallel camera model eliminates the keystone distortion and is able to create geometrically correct stereoscopic images [Woods et al., 1993]. However, shooting/rendering with parallel camera configuration is not always advisable or even practical for certain stereoscopic scenes [Ronfard and Taubin, 2010]. Consider

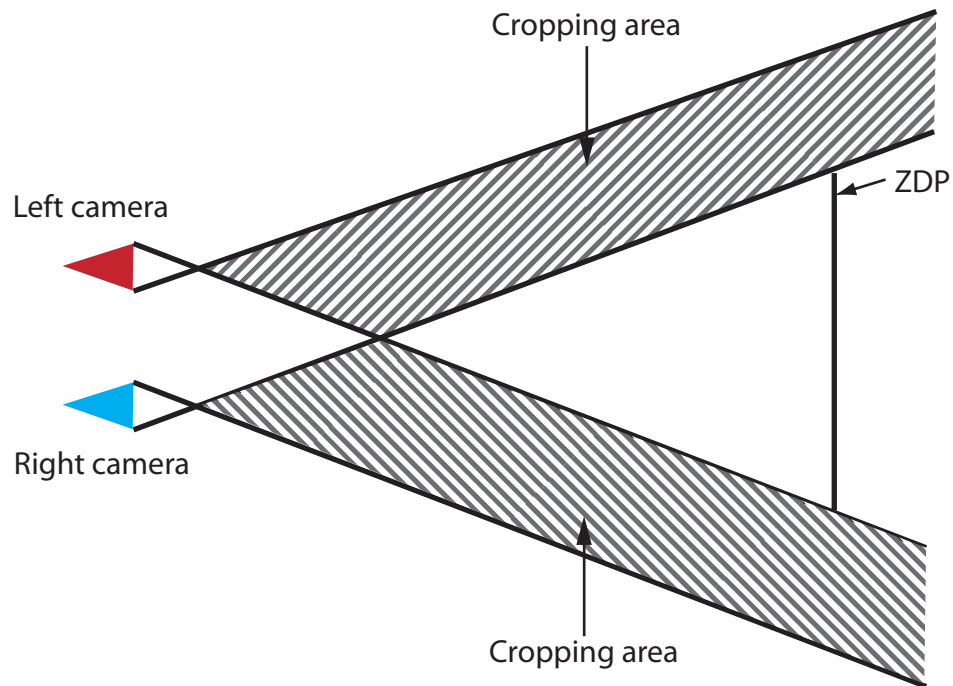


Figure 3.13: Symmetric parallel camera configuration [Froner and Holliman, 2005]

a scene where the filming subject is extremely close to the camera, the converged camera configuration is imposed otherwise the subject appears too close to the border of the viewing screen, producing the undesired “Stereoscopic Window Violation” (described later in this chapter). Nonetheless, stereoscopic cinematographers should always try to avoid the converged camera model and use the parallel camera arrangement to shoot/create S3D scenes.

3.2.3 Ghosting and Crosstalk

The crosstalk is an artifact where polarising filters allow a small amount of unwanted light to pass in their crossed state resulting in one eye seeing the remainder of the image intended for the other eye. Ghosting is the visible crosstalk which has the similar appearance to a double-exposed image. Ghosting is most likely to appear in high-contrast and high-screen-disparity (objects coming far inside/off the screen) 3D images. Also, viewers’ head movement (tipping) could substantially increase the level of crosstalk and easily induce ghosting. Lipton believes that a little bit of ghosting is as bad as the Vergence-Accommodation Conflict in terms of causing 3D

visual discomfort [Lipton, 2001]. In order to minimize this undesired artifact, the screen disparity value should be kept as low as possible for achieving the desired stereoscopic effect and high-contrast images and head movement should be avoided as much as possible.

3.2.4 Stereoscopic Window Violation

Consider an object that is behind the 3D screen, when it moves sideways and interferes with the left frame of the screen, a viewer sees more of the object with the right eye than the left as the sight of the left eye is blocked by the left frame of the screen. This is consistent with the real life viewing experience and does not introduce any inconsistent depth cues. However, when that object moves toward the viewer, crosses the 3D screen and stay on top of the left edge of the viewing screen. The viewer's left eye sees more of the object than the right eye does. Because the object is in front of the screen, the left image is shifted to the right revealing a portion of the object that is not visible to the right eye. The brain then faces a serious perceptual conflict between two depth cues: the eyes see an object with negative disparity as being in front of the screen, but the screen also appears to be in front of the object as it is occluding the object. The occlusion cue is stronger than the screen disparity and decides to push the object back to the screen. This undesirable effect is called the *Stereoscopic Window Violation*, SWV.

The SWV causes eyestrain or headache for most of viewers and should be avoided as much as possible [Lipton, 1982]. An object that is in front of the screen interfering with the top and bottom frames instead of the left and right frames of the screen are much easier for the brain to handle. The brain still has to deal with the inconsistent cues of the screen occlusion and the position of the object, but there is no difference in how much of the object is perceived in left and right eyes so that one source of the conflict is removed. The brain's likely solution is to bend the screen toward the viewer [Mendiburu, 2009].

The easiest solution to the SWV is to simply move the stereoscopic objects away from the borders of the screen. If that is not possible a virtual floating window, or the so-called "*Proscenium Arch*" can be created to be placed between cameras and

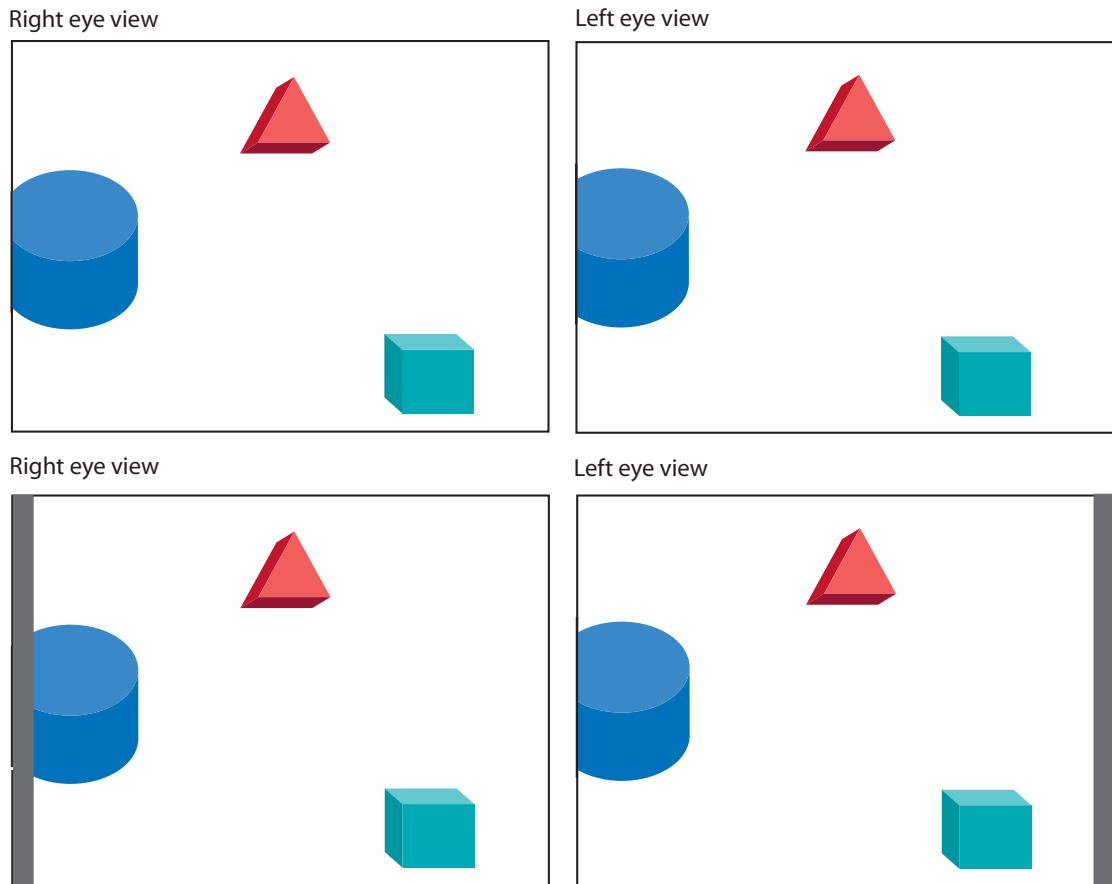


Figure 3.14: A Floating Window is created by blanking a small portion of the left side on the right image and right side on the left image [Autodesk, 2008]

the scene by blanking a small portion of each image (left side on the right image and right side on the left image), as illustrated in Figure 3.14. This process can be regarded as looking at the scene through a rectangle window which is inside a black wall. The black wall blocks parts of the images that interfere with the 3D screen. Note that the left and right images are swapped in Figure 3.14 because objects are in front of the screen and have negative screen disparity. Also, dark gray instead of solid black stripes are used to mask the images to reduce crosstalk. Other methods such as blurring or attenuating the intensity of the edges may also be tried [Ronfard and Taubin, 2010].

The SWV can be tolerated if an object is entering or exiting the frame rapidly, e.g., within half a second [Mendiburu, 2009]. The SWV does not produce any adverse effect if the whole object moves in/out of the frame before the brain can

perceived it as being in front of the screen. Also, images of objects with negative screen disparity, such as people, can be more easily forgiven if they are in dark colour or cast in shadow [Lipton, 1982].

3.2.5 Perceptual Distortions and Multi-rigging Technique

Two of the most discussed perceptual distortions in viewing S3D content are the *Cardboard Effect* and *Puppet-Theatre Effect* [Yamanoue et al., 2000, Masaoka et al., 2006].

Cardboard Effect and Puppet-Theatre Effect

The *Cardboard Effect* refers to the effect that a distant 3D object loses its roundness and appears flat to viewers. Considering an object that is hundreds of feet away from the viewer, rays coming from the object to the viewer's left and right eyes are essentially parallel. The brain can not distinguish any discrepancy between left and right images and no depth is perceived. Cinematographers can use the *Hyperstereo* technique to reduce this undesired effect. *Hyperstereo* refers to the process of capturing stereoscopic images with a camera separation larger than the human eye separation. By moving the left and right cameras further apart, rays from the distant object are coming into left and right eyes at different angles and depth perception is restored. This process results in a smaller stereoscopic scene, giving the viewer an impression of viewing through the eyes of a giant. In contrary to *Hyperstereo*, *Hypostereo* refers to the process of capturing the scene with a camera separation smaller than the human eye separation, creating an impression of viewing through the eyes of a very small creature.

An obvious tradeoff of the Hyperstereo technique is that the object in the foreground appears to be unnaturally small when the roundness of the object in the background is restored, which is commonly known as the *Puppet-Theatre Effect* [Yamanoue et al., 2006]. Both the Cardboard and Puppet-Theatre effects spoil the sense of reality and create perceptual distortions when viewing S3D content. Stereoscopic cinematographers developed the *Multi-rigging* technique to correct those perceptual distortions.

Multi-rigging Technique

The *Multi-rigging* technique refers to the process of capturing objects in foreground and background with different camera separations and composing foreground and background images together as illustrated in Figure 3.15. This technique was intensively used in S3D movies such as *Meet the Robinsons*, 2007 and *Beowulf*, 2007 [Engle, 2008].

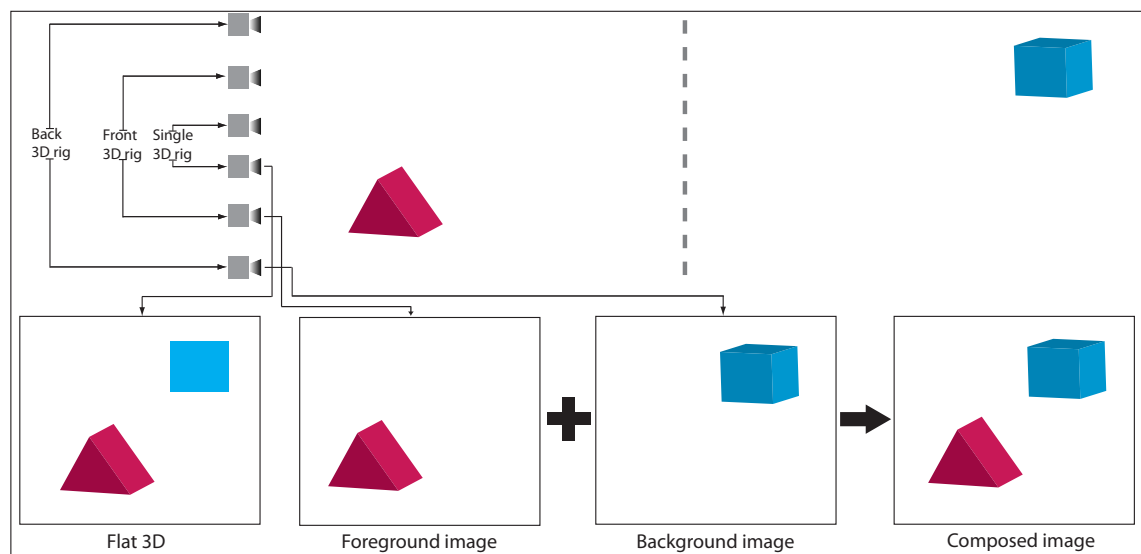


Figure 3.15: The Multi-rigging technique uses different camera separations to capture objects at different distances, eliminating the appearance of the Cardboard Effect and Puppet-Theatre Effect

Consider Figure 3.15, the triangle is close to the cameras and the cube is far away. In a single 3D camera rig, the cube appears to be flat due to the large distance to the cameras. In the Multi-rigging camera configuration, the roundness of the cube is able to be restored with a large camera separation. There is no Puppet-Theatre effect on the triangle as it is captured by a relatively smaller camera separation.

3.3 Cinematic Storytelling in Stereoscopic Cinema

The future of stereoscopic medium lays upon cinematographers, who have always been interested in adding the third dimension to the flat picture. Over the years cinematographers have developed certain so-called “*Cinematic Storytelling*” techniques, such as moving objects in space, camera motion, depth of field, lighting, and the use of fog and mist, to help make the single-eye viewed movies look three-dimensional.

However, the employment of cinematic storytelling techniques in S3D imposes a new challenge that cinematographers have never dealt with: the perceived depth keeps changing with the scene depth change. Careful consideration and a great deal of effort is required to ensure that the perceived depth remains inside the 3D Comfort Zone around the screen and there is no viewing discomfort throughout the whole presentation.

In order to thoroughly comprehend this problem, one must start with a basic understanding of the standard cinematic storytelling techniques and their applications in cinema.

3.3.1 Cinematic Storytelling Techniques

The term *Cinematic Storytelling* techniques refers to the non-dialog techniques used to convey ideas in movies. Before the use of synchronised sound and colour, cinematic storytelling techniques were the only way available to tell stories in early movies like *The Great Train Robbery*, *Metropolis* and *The Battleship Potemkin*. The non-dialog techniques such as object movement, camera motion, frame composition, and editing manipulate audiences’ emotions, carrying characters, and revealing plots without the audience’s immediate knowledge. For example, consider Steven Spielberg’s famous movie: *ET*, the first ten-minute set-up is completely cinematic and does not have a word of dialog or voice narration. However, a nine year-old is able to understand who the bad guy is and why. That is the reason that cinematic storytelling is so effective, engaging and frequently used.

As cinematic storytelling often operates on our subconscious, viewers are more

Cinematic Techniques	Method	2D Movie Example	Dramatic Value
Object Movement			
Screen Direction	Moving objects along the X/Y axis	<i>Strangers on a Train</i> (1951) <i>The Piano</i> (1993)	Exploring the spatial dynamics inherent in the film frame
Z-Axis and Depth of Field	Objects moving from foreground <-> background and going in/out of focus	<i>Citizen Kane</i> (1941) <i>The Graduate</i> (1967) <i>Dolores Claiborne</i> (1995)	Creating the illusion of depth and dividing the static frame into different planes
Camera Motion			
Tracking Shot	Moving camera along a track/path	<i>Marathon Man</i> (1976) <i>The 400 Blows</i> (1959) <i>Fatal Attraction</i> (1987)	Depicting the information in the “moving close-up”
Pan	Panning camera sideways	<i>Dances with Wolves</i> (1990)	Depicting the information in real time and suggesting a certain fluidity
Tilt-Up/Down	Tilting camera up/down	<i>Fargo</i> (1996) <i>Léon</i> (1994)	Revealing details that audiences may not notice otherwise

Table 3.2: Cinematic Storytelling Techniques

used to certain cinematic patterns which repeatedly appear in movies. For example, objects moving down the screen appear more natural to viewers than objects moving up as it is assumed that they are assisted by gravity. Therefore, if viewers see an object moving up the screen, they will automatically pay more attention to it. This is the reason that stereoscopic filmmakers should be fluent in using cinematic storytelling techniques in order to efficiently create a compelling and convincing viewing experience. In the book “*Cinematic storytelling: the 100 most powerful film conventions every filmmaker must know*” [Sijll, 2005], Jennifer Van Sijll assembled 100 cinematic storytelling techniques that are commonly used in current movie productions. In this study, we mainly focus on the following two techniques that can cause significant perceived depth change in stereoscopic movies.

Objects Movement

As demonstrated in Table 3.2, we name the first category: Object Movement standing for moving objects inside the scene volume. Screen Direction refers to the direction an object is travelling. Objects moving at different directions can imply

different meanings in movies such as antagonism, individualism, and conflict. Here we could only briefly discuss it as background information.

X and Y Axes

X-Axis refers to the axis that horizontally cuts the screen. An object can run left-to-right or right-to-left along the X-axis. Our eyes move more comfortably from left to right as it is consistent with the reading habit. Human eyes are less experienced in moving from right to left, therefore it is less comfortable. This is the reason that the protagonist normally enters the screen from the left in most of the movies. When the “good guy” moves along the X-axis left-to-right the audience’s eyes move comfortably, the audience subconsciously starts making positive interpretation. On the contrary to the protagonist, the antagonist normally enters the screen from the right, which makes the viewer uncomfortable as human eyes are not used to moving from right to left. Our learned discomfort is transferred to the “bad guy” by the screenwriter through the screen direction. This subtle irritant can be exploited to direct the audience to negatively perceive a character. The screen direction can be coded negatively in the same way that a black hat is coded as a symbol of negativity. When those two forces, the “good guy” and “bad guy”, are aiming at each other, the audience naturally expects some kind of a collision. This method was excellently adopted in the opening scene of Alfred Hitchcock’s *Strangers on a Train*.

Y-Axis refers to the line that runs along the screen vertically. An object can move up toward the top of the screen or down toward the bottom of the screen along the Y-axis. As mentioned above, moving objects down the screen seems to be easy as it is assumed that it is assisted by the gravity. Moving objects up the screen, on the contrary, seems to be difficult as it appears to be resisted by the gravity. When the audience sees an object move along the Y-axis in a linear established route with a fixed speed, he/she naturally assumes a “good” destination is ahead somewhere along the line. A linear established route along the Y-axis can represent normalcy and safety, whereas its opposite can also be established by a dangerous detour along the Y-axis. Such an example can also be found in *Strangers on a Train*.

Z-Axis and Depth of Field

Z-Axis refers to the axis that runs from the foreground to the background or background to the foreground in the screen. *Depth of Field* refers to the in-focus distance or the focus range along the Z-axis. The combination of Z-axis and depth of field gives the audience the sense of the 3D space and can be manipulated differently to convey different ideas. In *Citizen Kane*, Kane's inner turmoil was communicated by the change of his size when he walked back and forth along the Z axis. An extended and deep depth of field, which allows objects in the foreground and background to be in focus at the same time, was employed so that the figure of Kane was always sharp and clear. In *The Graduate*, a limited and shallow depth of field was created. Only a short distance along the Z-axis can be in focus at one time allowing filmmakers to selectively shift the audience's attention from one object to another anytime during the scene. In *Dolores Claiborne*, a shallow depth of field was used to divide the static frame into three separate zones along the Z-axis: foreground, middleground, and background, creating different time zones inside which different stories can be staged.

Camera Motion

Tracking Shot

A *Tracking Shot* takes place when the camera is mounted on a dolly moving smoothly along tracks, which can be straight or curved patterns. Tracking shot can be useful for various situations. In the opening scene of *American Beauty*, the camera tracked along the jury in the "moving close-up". In the opening diner scene of *Reservoir Dogs*, the camera tracked around the jewel thieves so that the audience can study each individual's face. The camera can also track along an object as it did in *Marathon Man* and *The 400 Blows*, in which the camera run alongside the protagonist.

Pan

A *Pan* takes place when the camera moves from left to right or right to left along a horizontal line. The movement of the camera results in the reveal of new information

which is not seen by the audience before. Another function of the pan shot is that it can depict information in real time with continuity. A Point of View (POV) shot is one application of the pan shot among many others. In the opening scene in *Dances with Wolves*, a pan shot revealed the protagonist's point-of-view as he scanned across the table and saw those bloody surgery tools. Given the information about what the protagonist saw, the audience can understand the reason that he decided not to perform the amputation.

Tilt-up\down

A *Tilt-up\down* occurs when the camera moves up or down along a vertical line. Much like the pan shot, the tilt-up\down shot is often used to reveal new information. It could be an important clue, a hidden character, or the parameters of a location. The character of young Mathilda was first introduced to the audience by a tilt-up shot in *The Professional*. The camera slowly tilted up from her boots to her face, giving the audience time to study the contradictions in her character. In *Fargo* there is an example of a conventional use of the tilt-down shot. The shot started with a freeway sign. As the camera tilted down, a car was revealed, establishing the location of the car.

3.3.2 Cinematic Storytelling in S3D

Employing cinematic storytelling techniques often involves changing the scene depth that is the volume or boundaries of all the objects inside the scene. The scene depth change makes no difference in conventional 2D movie as everything appears to be flat on the screen. However, stereoscopic content produce perceived depth on the display which changes proportionally with the magnitude of scene depth.

Perceived Depth Change in S3D

In stereoscopic movies, objects moving along the Z-axis is one of the main factors that causes perceived depth change. As shown in Figure 3.16, a spaceship is flying from point **A** to **C** along the flight path. The world disparity, d_B , becomes tiny when the spaceship comes close to the Zero Disparity Plane, ZDP (point **B**) and increases

as the spaceship flies away from the ZDP, d_A and d_C . As the world disparity is to be mapped to screen disparity in the same proportion, the perceived depth on the display also changes with the movement of the spaceship and could be out of the limits of the 3D Comfort Zone.

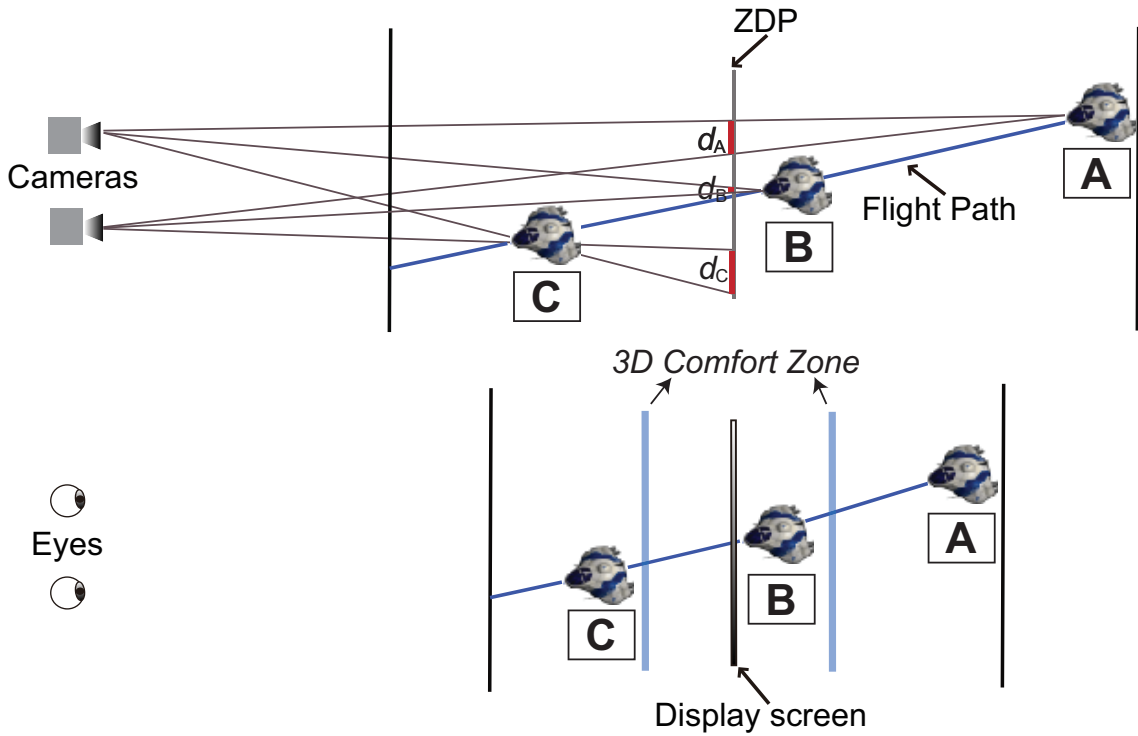


Figure 3.16: Perceived depth change caused by object movement. When the spaceship flies from A to C, the perceived depth on the display also changes with the movement of the spaceship and could be out of the limits of the 3D Comfort Zone.

Perceived depth changes can also be imposed by Camera Motion. Pushing camera in and out of an object effectively moves the object along the Z-axis in the scene; panning camera sideways or tilting camera up and down can reveal new objects which are not exposed before. The perceived stereoscopic depth then changes according to the pattern of scene depth change. Regarding Figure 3.17, suppose the camera separation and the distance between cameras and the ZDP stay constant as the cameras move along the predefined path which runs from A to C. The world disparity, d_A is much larger than d_B as the spaceship at A is quite far away from the ZDP and the spaceship lies upon the ZDP at B. Similarly, the perceived depth on the screen changes dramatically as the cameras move from B to C.

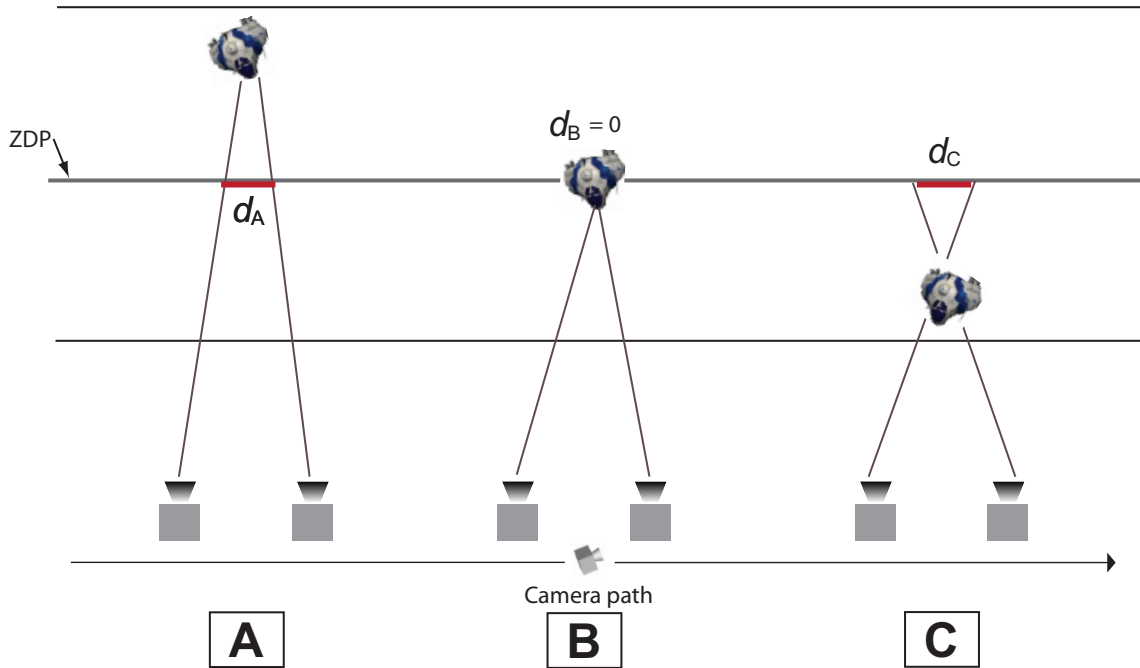


Figure 3.17: Perceived Depth change caused by camera motion. The perceived depth on the display screen changes dramatically as the cameras move from A to B or B to C.

We review literature on how to efficiently control the perceived depth in S3D cinematography in Chapter 4.

3.4 Summary

In this chapter, we provided a historical review of the cinema technology from the invention of sound and colour to the stereoscopic cinema, followed by a detailed discussion about key factors on producing good quality stereoscopic movies: the perceived stereoscopic depth needs to be limited to ensure the viewing comfort; parallel camera configuration is required to capture stereoscopic images without vertical disparity caused by the keystone distortion; high-disparity and high-contrast stereoscopic content should be avoided as much as possible to reduce the ghosting and crosstalk; the Stereoscopic Window Violation needs to be eliminated especially on the vertical (left and right) borders of the 3D screen to prevent inconsistent depth cues; the Multi-rigging technique can be applied to correct the *Cardboard* and

Puppet-Theatre effects and enhance the 3D viewing experience. Those factors have already been well understood and efficiently manipulated in practice by stereoscopic cinematographers.

New research challenges have been brought by employing cinematic storytelling techniques in S3D cinematography. The cinematic storytelling techniques, such as the object movement, camera motion, and zooming, can alter the scene depth resulting in the dynamic change of perceived depth within and between a series of frames (shots). Literature on how to efficiently control the perceived depth in S3D cinematography will be discussed in the following chapter.

Chapter 4

Previous Works on Perceived Depth Control in S3D

Preceding chapters described background information on the human visual system, the design and application of different stereoscopic display systems, and stereoscopic cinematography. In this chapter, we discuss previous studies that are specifically related to our research problem: how to efficiently control the perceived depth to reduce the viewing discomfort imposed by the Vergence-Accommodation (VA) Conflict. The related literature is categorised into the following three areas: (1) identifying the limits of the 3D Comfort Zone, (2) mapping the scene depth to the perceived depth, and (3) applying the Depth of Field (DoF) blur simulation on stereoscopic content.

4.1 Limits of the 3D Comfort Zone

There are two general rules regarding the limits of the 3D Comfort Zone in literature:

- **Percentage of Screen Width Rule.** The crossed screen disparity, introducing perceived depth in front of the display, should not exceed 2-3% of the width of the display screen; the uncrossed screen disparity, introducing perceived depth behind the display, should not exceed 1-2% of the width of the display screen [Mendiburu, 2009].

- **No appearance of eye divergence.** The perceived stereoscopic depth should not cause the eyes of the viewer to diverge, i.e., the screen disparity (both crossed and uncrossed) should not be larger than the human eye separation. It is a consensus in literature that the eye divergence causes serious S3D viewing discomfort in form of eye strain or fatigue for viewers [Spottiswoode et al., 1952, Lipton, 1982, McAllister, 1993, Mendiburu, 2009, Ronfard and Taubin, 2010].

In addition to these two general rules, several human factors studies have been conducted in defining the specific limits of the 3D Comfort Zone on a chosen 3D display.

4.1.1 Williams and Parrish's Results

Experiment Method

Williams and Parrish performed an experiment with four subjects, where each subject was asked to move a real-world probe to the position of a virtual 3D rod four times on each perceived depth plane tested. This process was repeated in a wide range of perceived depth planes. The information about the exact range of perceived depth tested and experiment apparatus was not provided in their paper [Williams and Parrish, 1990].

Experiment Results and Limitations

For each perceived depth plane tested, there were 16 results ($4 \text{ subjects} \times 4 \text{ times}$) of the position of the probe. The authors believed that the position of the rod should fall outside the 95% confidence interval of the position of the probe when the corresponding perceived depth of the rod exceeded the 3D Comfort Zone, and *vice versa*.

Three viewing distances: 483 mm, 965 mm, and 1448 mm were tested and the results of the limits of the 3D Comfort Zone varied with different viewing distances as illustrated in Figure 4.1. Based on the observation of the results, the authors concluded that the limits of the 3D Comfort Zone should be: 25% of the viewing

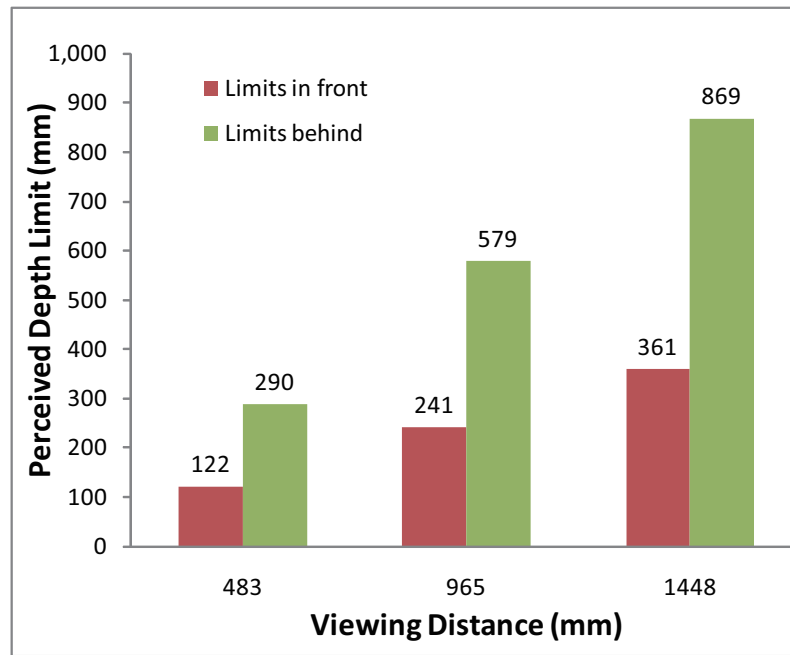


Figure 4.1: Williams and Parrish's results on the limits of the 3D Comfort Zone

distance in front of the display and 60% of the viewing distance behind the display. In addition they also asserted that increasing the viewing distance could expand the limits of the 3D Comfort Zone [Williams and Parrish, 1990].

The limitation of this work was that it could not efficiently distinguish the excessive perceived depth. Consider the situation where the perceived depth was excessive and subjects were struggling to fuse the rod on the display, the results of the position of the probe among subjects should have a large variation, resulting in a large 95% confidence interval, which could still include the position of the rod. Therefore, this excessive perceived depth would be deemed as being inside the 3D Comfort Zone.

4.1.2 Yeh and Silverstein's Study

Experiment Setup

Ten subjects were asked to fuse red or white solid-fill T-shaped images of the size of 7.01 mm (horizontal) \times 7.13 mm (vertical). The T-shaped image was horizontally centered at one of the three possible vertical screen positions: top, middle, and bottom. The 3D display used was a 16" Time-Multiplexed CRT 3D display (Tektronix SGS-430 Stereoscopic Colour Graphics System). The viewing distance was 660 mm.

Experiment Procedure

For each colour (red or white) and vertical position of the image (top, middle, bottom), the value of the screen disparity increased/decreased from a randomly selected starting point until the subject was neutral about whether he/she could fuse the displayed T-shape (two separate images or one single 3D image), and the resulting screen disparity value was taken as the threshold for this subject.

Experiment Conclusion and Limitations

The limit of the screen disparity value was obtained by averaging the threshold values among all ten subjects for both image colours and all three vertical screen positions, which was 27.0 min arc for the crossed screen disparity (csd) and 24.0 min arc for the uncrossed screen disparity (usd). Given the viewing distance of 660 mm, the screen disparities in millimeter can be derived as follows:

$$csd = 660 * \tan(27/60) = 5.2 \quad \text{and} \quad usd = 660 * \tan(24/60) = 4.6 \quad (4.1)$$

With the nominal eye separation of 65 mm [Ronfard and Taubin, 2010], the perceived depth, p , can be obtained using:

$$p_{front} = \frac{z}{\frac{e}{csd} + 1} = 49 \quad \text{and} \quad p_{behind} = \frac{z}{\frac{e}{usd} - 1} = 50 \quad (4.2)$$

The limits of the 3D Comfort Zone defined by Yeh and Silverstein were 49 mm in front of the display and 50 mm behind the display with a viewing distance of 660 mm [Yeh and Silverstein, 1990].

This work suffered major drawbacks in that its results on the limits of the 3D Comfort Zone was obtained by averaging the threshold values of the perceived depth limit of all subjects. The threshold value was found based on the individual viewer's subjective judgement on whether he/she could fuse the tested stimuli, which could result in a large variance of the threshold values among subjects. As the individual threshold values were not even provided or discussed, it was difficult to believe that the defined limits of the 3D Comfort Zone were applicable.

4.1.3 Woods *et al*'s Experiment

Experiment Apparatus and Subjects

A 100Hz Time-Multiplexed 3D display was employed in conjunction with a 16" monitor and Tektronix polarising screen. 10 subjects were recruited. The viewing distance was approximately 800 mm.

Experiment Method

For each subject, a 3D image of a donut (40 mm in diameter) was displayed with increased or decreased screen disparities. The increase of the screen disparity began with placing the donut on the display screen and gradually increasing the value of the screen disparity (crossed or uncrossed) until the subject could not fuse the 3D image. The decrease of the screen disparity started by displaying the 3D donut with the screen disparity value equal to the screen width and gradually reducing the value of screen disparity until the subject could fuse the 3D image. Each subject saw the 3D image on each tested perceived depth plane at least three times.

Experiment Results

This work did not conclude a range for the 3D Comfort Zone due to a large variation among individual results [Woods et al., 1993]. However, the maximum screen disparity value (both crossed and uncrossed) that all subjects could fuse was 10 mm, equivalent to 107 mm perceived depth in front of the display and 145 mm perceived depth behind the display with a viewing distance of 800 mm.

4.1.4 Jones *et al*'s Results

Experiment Apparatus and Subjects

A 13.8" auto-stereoscopic display with a LCD screen was used. Eight subjects were tested with a viewing distance of 700 mm.

Experiment Method

Two test stimulus were included: a 3D textured box and a 3D Mallet Box [Tang and Evans, 2007]. For each stimulus, the subject was allowed to alter the 3D object's perceived depth both in front of and behind the display at will until the fusion of the 3D image was lost. The subject was required to look away from the display before trying to fuse the 3D image on another perceived depth plane.

Experiment Results

The results under two stimulus were significantly different:

With the simple stimuli of the textured box, the perceived depth limit in front of the display was between 200 mm and 400 mm, the perceived depth limit behind the display was found between 500 mm and 1000 mm.

With the more sensitive stimuli of the Mallet box, the perceived depth limit in front of the display was found between 50 mm and 190 mm, the perceived depth limit behind the display was found between 60 mm and 180 mm.

This work also suffered from a large variation among individual results. The authors concluded a range for the 3D Comfort Zone based on the smallest perceived depth that could be fused by all 8 subjects, which was 50 mm in front and 60 mm behind the display with a viewing distance of 700 mm [Jones et al., 2001].

4.2 Existing Depth Mapping Algorithms

Before the discussion of existing depth mapping approaches, we introduce the following two terms: static scene and dynamic scene.

The term “Static Scene” here refers to a scene that has a constant scene depth value. Such a static scene is commonly seen in 2D/3D applications. It can be a scene in which all objects are static or a scene in which certain objects are moving but the dynamic objects never move outside the boundaries set by static objects. Consider a scene demonstrated in Figure 4.2(a), the spaceship flies around the still asteroids yet never exceeds the boundaries set by those asteroids, therefore, the value of the scene depth remains the same.

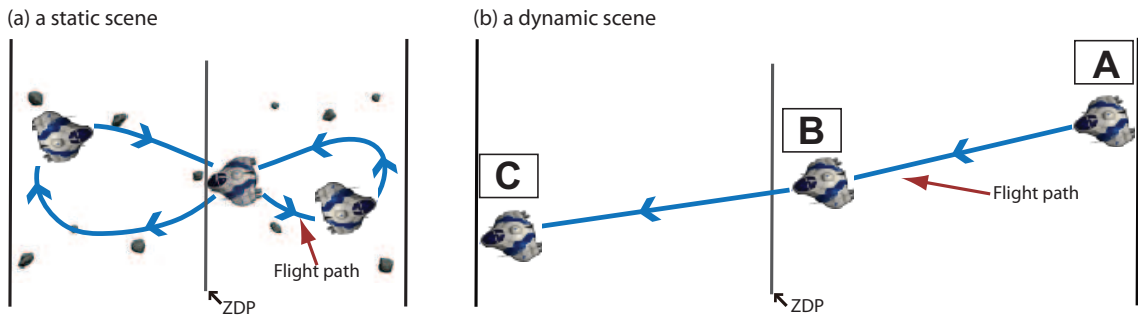


Figure 4.2: An example of a static scene and dynamic scene

In contrast to a static scene, a “Dynamic Scene” in this study refers to a scene whose scene depth value changes over time due to the employment of cinematic storytelling techniques such as the object movement, camera motion, etc. Consider Figure 4.2(b), there is only one scene object: a spaceship flying from A to C. When the spaceship flies from A to B, the scene depth decreases as B is closer to the ZDP than A; when the spaceship flies from B to C, the scene depth increases as C is further away from the ZDP than B. The scene depth value is decided by and changes with the position of the spaceship.

Note that our definition of a static scene is different from the traditional cinematography term “lock-down” or “lock-off” shot in which the camera is locked down on a tripod [Sijll, 2005]. A static scene and dynamic scene is separated by the scene depth change not by the movement of the camera or scene object.

A few depth mapping algorithms have been proposed to control the mapping from scene depth to perceived depth in a static stereoscopic scene.

4.2.1 Ware *et al*’s Algorithm

Ware *et al*’s algorithm incorporated the results of a human-based experiment in which viewers were given control over the virtual camera separation. The experiment showed that viewers were inclined to increase virtual camera separation in flat scenes and decrease virtual camera separation in deep scenes. This information was used to calculate a so-called “Cyclopean scale” which could guarantee the nearest part of the scene came to be just behind the screen as shown in Figure 4.3. The effect of this scaling was that the size of the image remained unchanged on the display

while the screen disparities increased for distant objects and decreased for close objects. The cyclopean scale was then used to derive an equation that calculated the camera separation as a function of the scene depth. A damping factor was applied to account for individual changes of the virtual camera separation [Ware et al., 1995, Ware et al., 1998].

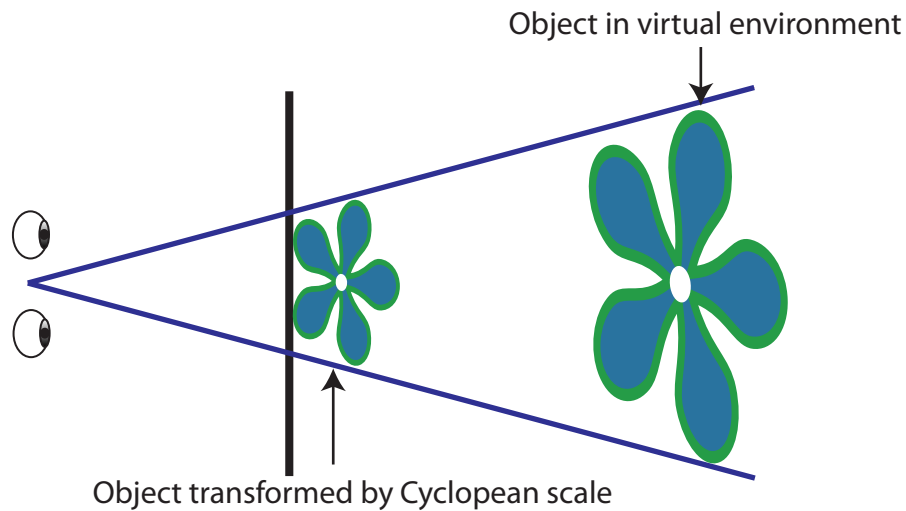


Figure 4.3: The cyclopean scale [Ware et al., 1995, Ware et al., 1998] which can guarantee the nearest part of the scene comes to be just behind the screen.

This algorithm was evaluated by displaying scene objects behind the display. For flat scenes with less than 30 cm scene depth, the perceived depth was doubled; for deep scenes with a scene depth larger than 70 cm, the perceived depth was limited.

Limitations

This algorithm cannot precisely control the stereoscopic depth perceived on display as there was no direct mapping between scene depth and perceived depth. It simply doubled the perceived depth for small scene depth values and limited the perceived depth for large scene depth values. The S3D viewing discomfort can still easily appear because of dramatic scene depth changes or excessive perceived depth as it allowed the user to alter the virtual camera separation.

4.2.2 Wartell's Algorithm

Figure 4.4 illustrates the distortion caused by false-eye modelling (underestimating human eye separation) when viewing stereoscopic images from different directions on a head-tracked stereoscopic display. Two sets of real eyes are illustrated in large black circles and inside each set of real eyes there is a set of modelled eyes illustrated in black squares. For each set of real eyes, the modelled point (large grey circle) is projected onto the 3D display by modelled eyes. The real eyes perceive the modelled point by finding the intersection of black lines which are constructed by connecting an real eye with its corresponding projected image, illustrated in grey squares, on the display. Note that the perceived point, illustrated in grey squares, moves with the eye movement and is closer to the screen than the modelled point.

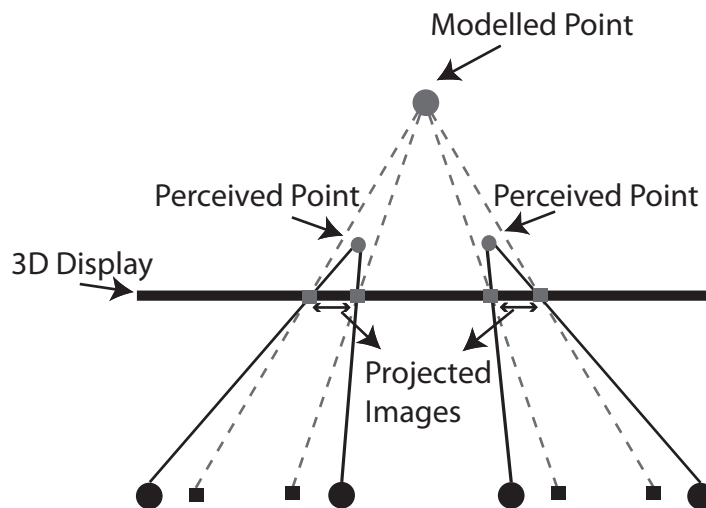


Figure 4.4: A geometric construction of the distortion caused by using false-eye modelling [Wartell et al., 1999].

Wartell [Wartell, 2001] studied this distortion and proposed a transformation matrix to correct the distortion by pre-distorting the scene. The equation to calculate the camera separation was derived from the screen disparities of the furthest fusible distances (both in front of and behind the screen). This method was evaluated by displaying 3D images on a head-tracked 3D display. The results showed that the transformation matrix could correct the distortion caused by the false-eye modelling.

Limitations

This algorithm could not directly map the scene depth onto the perceived depth range either. The desired perceived depth range could only be achieved after repeated trials. In addition, this method could only be employed to generate computer graphics 3D images due to the use of the transformation matrix.

4.2.3 Jones *et al*'s and Holliman's Algorithms

Jones *et al* [Jones et al., 2001] developed an algorithm that could map a given scene depth range onto any defined single region perceived depth range automatically as illustrated in Figure 4.5(a). The near and far limits of the scene were identified and the camera separation was then automatically calculated removing the problem of choosing the stereo camera separation from users. Comparing with Ware *et al*'s and Wartell's algorithms, it did not require the user to adjust the camera separation (multiple times) to achieve the desired perceived depth range and instead it directly provided the precise perceived depth range specified by users. This approach was evaluated by displaying a 3D teapot on a head-tracked 3D display. The results confirmed that it could precisely control the perceived depth without introducing any scene distortion.

Holliman [Holliman, 2004, Holliman, 2005] further developed a new piecewise, so-called Region of Interest (ROI) algorithm, shown in Figure 4.5(b). This method was a piecewise approach allowing users to subjectively partition the scene depth volume, with freedom to allocate preferential stereoscopic depth to the region of interest; an approach that could be seen as zooming in depth. 4.5(b), the perceived depth for near and far regions are compressed allowing the majority of perceived stereoscopic depth to be given to the region of interest.

Jones *et al*'s and Holliman's algorithms can intuitively and precisely map the depth from the scene to display space and can be used to control the stereoscopic depth perception effectively in static stereoscopic scenes. Cinematographers only need to provide certain scene and display characteristics and the desired perceived depth range is achieved even with the option of the preferential stereoscopic depth

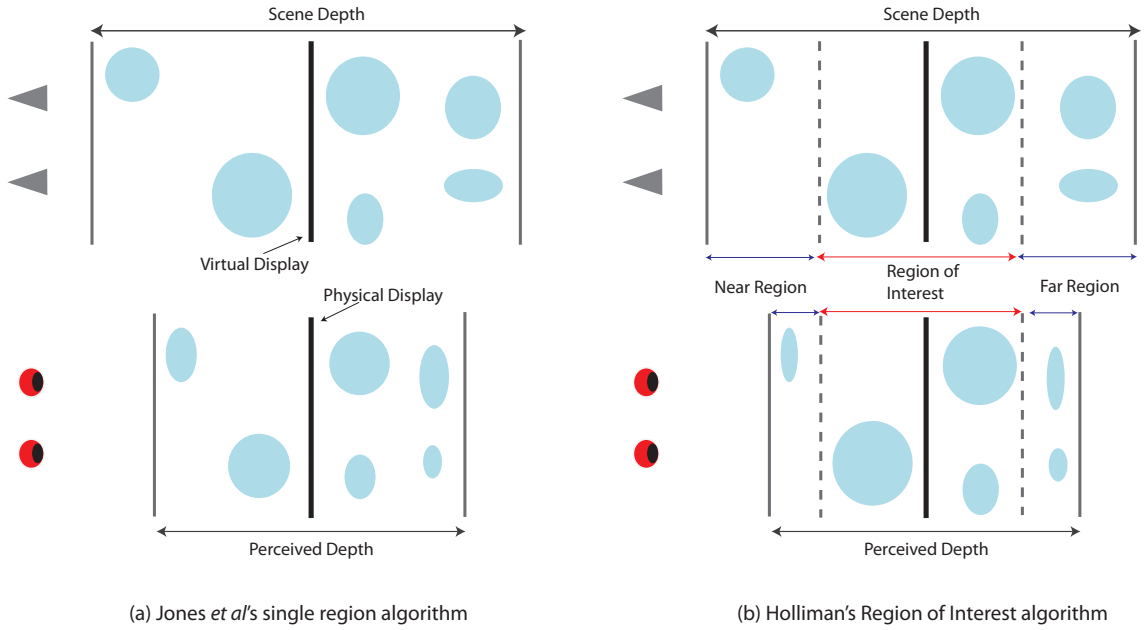


Figure 4.5: Jones *et al*'s and Holliman's Algorithms [Holliman, 2005]. (a) Jones *et al* mapped the scene depth as a whole. (b) Holliman's approach allowed users to subjectively partition the scene depth and allocate more perceived depth to the region of interest.

zooming in specified regions.

Limitations

Both algorithms used the so-called "Fixed Depth Mapping" method where the scene depth is mapped to the perceived depth in a fixed ratio by using a constant camera separation.

Consider Figure 4.6, the maximum scene depth range is mapped to the limits of the 3D Comfort Zone to ensure comfortable 3D viewing. When the spaceship is at A or C which has the maximum scene depth value, the perceived depth of the spaceship, p_a or p_c , reaches its maximum; when the spaceship is at B, the perceived depth of the spaceship, p_b , becomes very small as B is very close to the ZDP.

This fixed depth mapping method has the following two drawbacks when dealing with perceived depth control in dynamic stereoscopic scenes.

- The perceived depth can become inadequate when the scene depth is small.

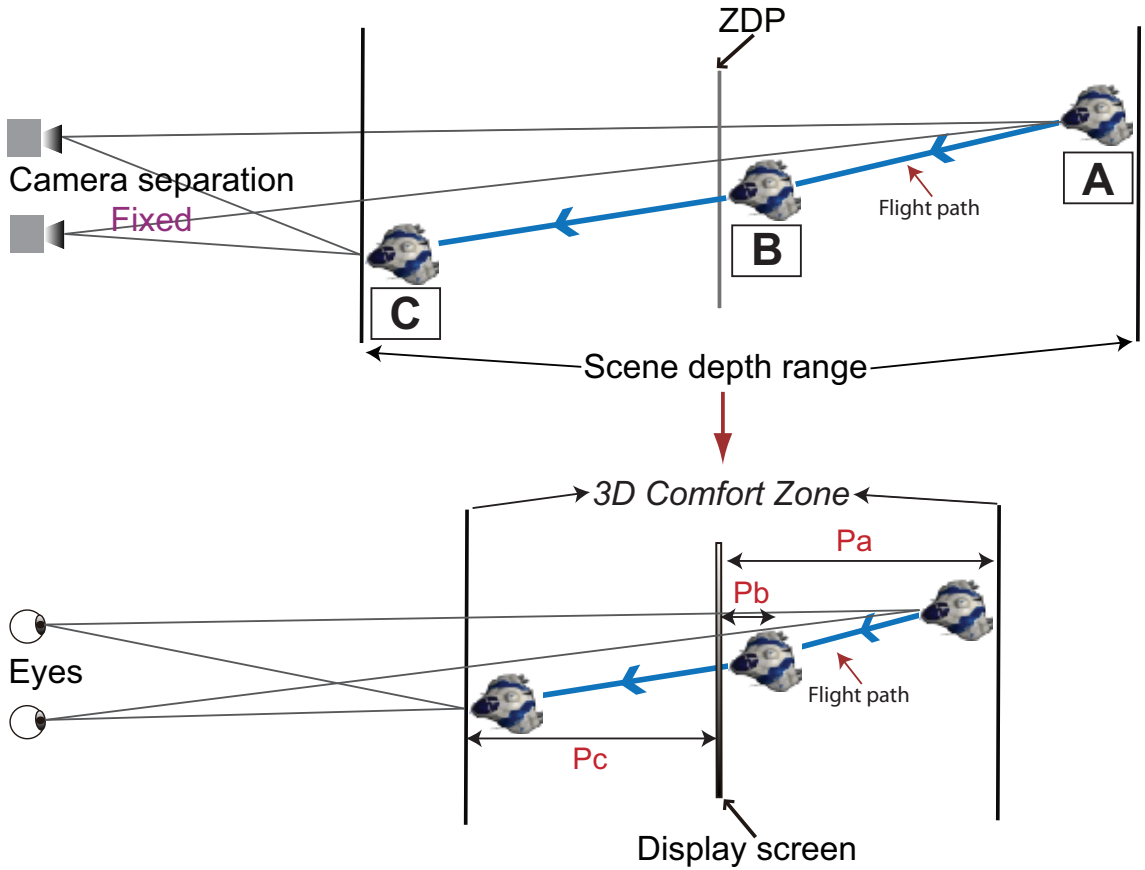


Figure 4.6: Jones *et al*'s algorithm in a dynamic stereoscopic scene

- Sudden and dramatic perceived depth change, which is a great contributor to the viewing discomfort [Mendiburu, 2009], can still take place when the scene depth changes suddenly and dramatically. Consider the situation where the spaceship flies promptly from A to B in Figure 4.6, the perceived depth changes suddenly and dramatically, $p_a \gg p_b$.

4.3 Depth of Field Blur Simulation in S3D

The purpose of applying Depth of Field (DoF) Blur effect or the so-called “retinal image blur” [Hoffman et al., 2008, Watt et al., 2005a] in cinematography is to simulate the real life viewing experience. When human eyes are looking at an object in the real world, the eyes are focusing on that object and creating a focus range around that object inside which objects are in focus and seen in full sharpness; objects outside the focus range are out of focus and appear blurred. The degree of blurring is

in proportion with the distance to the focus. The DoF of the human eye is between 0.2 D [Atchison and Smith, 2000] and 0.3 D [Campbell, 1957](Diopters, 1/metres), which can be slightly affected by parameters such as the pupil size, wavelength, and spectral composition [Marcos et al., 1999]. The Focus Range, FR, can be derived from:

$$FR_{front} = \frac{1}{\frac{1}{d_f} + DoF} \quad \text{and} \quad FR_{behind} = \frac{1}{\frac{1}{d_f} - DoF} \quad (4.3)$$

Given a focus distance (d_f) of 1m, a DoF of 0.2D, the focus range is approximately from 0.83m to 1.25m, i.e., when a viewer is looking at an object that is 1m away, objects with the distance ranging from 0.83m to 1.25m to the viewer are also in focus. It is obvious from the formula that the DoF is independent of the focus distance but the focus range is in proportion with the focus distance.

Research conducted by Yano *et al* suggested that the limits of the 3D Comfort Zone on the display should match the focus range around the 3D display, i.e, the perceived depth should be limited inside the focus range around the display screen [Yano et al., 2004].

4.3.1 Yano *et al*'s Experiment

Experiment Setup

A 120 Hz 28" Time-Multiplexed HDTV image monitor was used as the stereoscopic display. Six subjects were recruited. The viewing distance was 108 cm (105 cm was shown as the viewing distance in Figure 3 of their paper [Yano et al., 2004]).

Experiment Method

The subject was asked to read text for 64 minutes (three cycles of 15 min of reading and 3 minutes of rest, then another 10 min of reading: $(15 + 3) \times 3 + 10 = 64$). The text was displayed in four different perceived depth ranges: 0 (seeing the text on the display screen), 22.7 cm in front and 136.5 cm behind, 32.2 cm in front and 170.6 cm behind, and 39.7 cm in front of and 225.9 cm behind. The second perceived depth range (22.7 cm and 136.5 cm) matched the focus range around the display with a DoF of 0.2D under the given viewing distance. The last two perceived depth

ranges were designed to test the visual comfort with perceived depth exceeding the focus range around the display.

A subjective assessment, asking the subject to rate the degree of visual fatigue, was carried out after the experiment. An objective test was performed by measuring the subject's accommodation response (focus ability) before and after watching the tested stimuli.

Experiment Results

For each tested perceived depth range, the results of both objective and subjective assessments were averaged. The authors concluded that: (1) the perceived depth range that matched the focus range around the display (22.7 cm in front of and 136.5 cm behind), did not cause more significant visual discomfort than viewing the tested stimuli on the display screen (i.e., zero perceived depth). (2) Visual fatigue was caused by perceived depth that exceeded the focus range around the display. Therefore, the perceived depth should be limited inside the focus range around the display screen for visual comfort [Yano et al., 2004].

Limitations

It would be difficult to generalise their results on different 3D display systems due to the lack of statistical analysis (only mean scores of each perceived depth range tested were provided).

4.3.2 Blohm et al's Study

Blohm *et al* conducted a study on comparing the viewing experience with different ranges of DoF in stereoscopic scenes [Blohm et al., 1997].

Experiment Setup

Two identical 21 " monitors and a mirror stereoscope were used as the displaying system. Eight subjects (5 male and 3 female of the age between 25 and 35) were recruited. The tested stimuli was a 8 second stereoscopic video sequence captured

from a scene consisted of a textured background, several static objects placed on or above a static quadrilateral platform, and a flying ellipsoidal object. The perceived depth range tested was 17 cm in front and 32 cm behind the display. There were 6 different 3D video sequences, one without the DoF and 5 with the DoF of 5 different ranges (gradually increased from a range smaller than human's DoF to one larger than human's DoF). The DoF was always centered around the only moving scene object: the ellipsoidal object. The viewing distance was 70 cm.

Experiment Method and Results

Each subject watched 12 video sequences (each sequence twice in a random order). The subject was asked to always focus the ellipsoidal object which flew back and forth along a curvilinear path through objects on the platform. A subjective assessment was carried out by asking the subject to rate the viewing discomfort (1 being imperceptible,..., 5 being very annoying) after watching all 12 video sequences.

The results of the subjective assessment were analysed by the ANOVA for repeated measurement designs. The authors concluded that applying the DoF simulation could improve the S3D viewing experience, and a DoF that is slightly smaller than that of the human eye would actually be preferred in viewing stereoscopic scenes.

Limitations

In this study, the subjects were asked to keep their focus on the only flying scene object all the time, which was not consistent with the real-life 3D viewing experience. Research has confirmed that the eye movement of the viewers would be more widely spread in viewing S3D content [Hakkinen et al., 2010]. Whether viewers would still prefer the DoF blur simulation when they have the freedom to watch any object in the scene, introducing the possibility of viewing objects that are outside the DoF, remains to be an open question.

4.4 Summary

In this chapter, we reviewed literature on how to reduce the S3D viewing discomfort caused by the VA Conflict in the following three areas:

- **Limits of the 3D Comfort Zone.**

Both Williams and Parrish's and Yeh and Silverstein's studies [Williams and Parrish, 1990, Yeh and Silverstein, 1990] had serious flaws in their statistical analysis methods. Regarding research conducted by Wood *et al* [Woods et al., 1993] and Jones *et al* [Jones et al., 2001], no statistical method was discussed and the results of their experiments could not be generalised to determine a recommendation for the limits of the 3D Comfort Zone.

In this study, we aim to provide a new method that efficiently identifies the statistically meaningful limits of the 3D Comfort Zone on a given 3D display. This method will be described in Chapter 6.

- **Depth Mapping Methods.**

Both Ware et al's and Wartell's depth mapping methods [Ware et al., 1995, Ware et al., 1998, Wartell, 2001] can not effectively control the depth mapping from scene space to display space. Jones *et al*'s and Holliman's algorithms [Jones et al., 2001, Holliman, 2004, Holliman, 2004] can precisely map the scene depth range to a predefined perceived depth in static stereoscopic scenes. However, their fixed depth mapping method can result in inadequate perceived depth, and dramatic and sudden change of the perceived depth when dealing with dynamic stereoscopic scenes.

We developed a new dynamic depth mapping method that ensures the viewer's perceived depth stays constant on the 3D display and is always mapped to the limits of the 3D Comfort Zone, providing maximum perceived depth without causing the viewing discomfort. This method will be described in Chapter 5.

- **DoF Blur Simulation.**

Yano *et al*'s study indicated that the perceived depth should be limited inside the focus range around the display to ensure viewing comfort. However, the

lack of statistical analysis made it difficult to apply their results on other 3D display systems. Although Blohm *et al*'s experiment suggested the DoF could improve the visual comfort in viewing S3D motion pictures, the results of their experiment were biased by limiting the viewer's focus point, which contradicted the real-life S3D viewing experience.

In Chapter 5, we evaluate the DoF simulation in an experiment where the subject is free to focus on any object in the scene. The results of the DoF blur simulation are statistically compared with those of depth mapping methods to decide an ideally method for efficiently controlling the perceived depth in S3D.

Chapter 5

Investigation and Evaluation of Perceived Depth Control Methods

As discussed in the previous chapter, existing perceived depth control algorithms can efficiently control the depth perception in static stereoscopic scenes. However, current algorithms do not normally support standard Cinematic Storytelling techniques. These techniques, such as object movement, camera motion, and zooming, can result in dynamic scene depth change within and between a series of frames (shots) in stereoscopic cinematography. In this chapter, we empirically evaluate the following three types of perceived depth control methods that aim to address this problem.

(1) *Real-Eye Configuration*: set camera separation equal to the nominal human eye interpupillary distance. The perceived depth on the display is identical to the scene depth without any distortion. (2) *Mapping Algorithm*: map the scene depth to a predefined range on the display to avoid excessive perceived depth. A new method that dynamically adjusts the depth mapping from scene space to display space is presented in addition to an existing fixed depth mapping method. (3) *Depth of Field Simulation*: apply Depth of Field (DOF) blur effect to stereoscopic content. Only objects that are inside the DOF are viewed in full sharpness. Objects that are far away from the focus plane are blurred.

We performed a human-based experiment using the Recommendation ITU-R BT.500-11 [Union, 2002] to compare the quality of stereoscopic video sequences

generated by the above-mentioned perceived depth methods. Our results indicated that previously defined limits of the 3D Comfort Zone for static stereoscopic images may have become too conservative for stereoscopic motion pictures, and different displays differ in their individual limits of the 3D Comfort Zone. New methods that can efficiently define the limits of the 3D Comfort Zone for individual 3D display are needed in order to better control the perceived depth in stereoscopic cinematography. Our new dynamic depth mapping method does have an advantage over the fixed depth mapping method. The DOF blur effect does not improve the perceived depth quality in 3D cinematography as expected. We anticipate the results will be of particular interest to 3D filmmaking and real time computer games.

5.1 Investigation of Different Perceived Depth Control Methods

Speranza *et al* [Speranza et al., 2006] also investigated the relationship between perceived depth, object motion and viewing comfort using stereoscopic video sequences. 21 subjects were tested on four 21" CRT 3D displays with a viewing distance of 104 cm. The subject watched 3D objects with different perceived depth coming in and out of the display screen at different speeds.

Their results, analysed using a modified version of the Single Stimulus method from the ITU Recommendation [Union, 2002], suggested that the speed of perceived depth change might be more important than the absolute magnitude of the perceived depth in determining visual comfort. However, their work suffered from using the toed-in camera model to construct the test stimulus. The perceived depth range could not be precisely controlled by the toed-in camera model. In addition, the toed-in camera model is well known for creating *Vertical Disparity* which itself is a great contributor to the S3D viewing discomfort (explained in detail in section 3.2.2).

In this study, we investigated the following perceived depth control methods which are based on the parallel camera configuration.

5.1.1 Real-Eye Configuration

As illustrated in Figure 5.1, the camera separation is set equal to the nominal human eye interpupillary distance of 65 mm [Dodgson, 2004] with the goal to simply map the whole scene on top of the display. We expected that the 3D video sequence generated by this method would cause viewing discomfort for viewers as human factors studies have confirmed the need of compressing scene depth around display space in viewing stereoscopic materials.

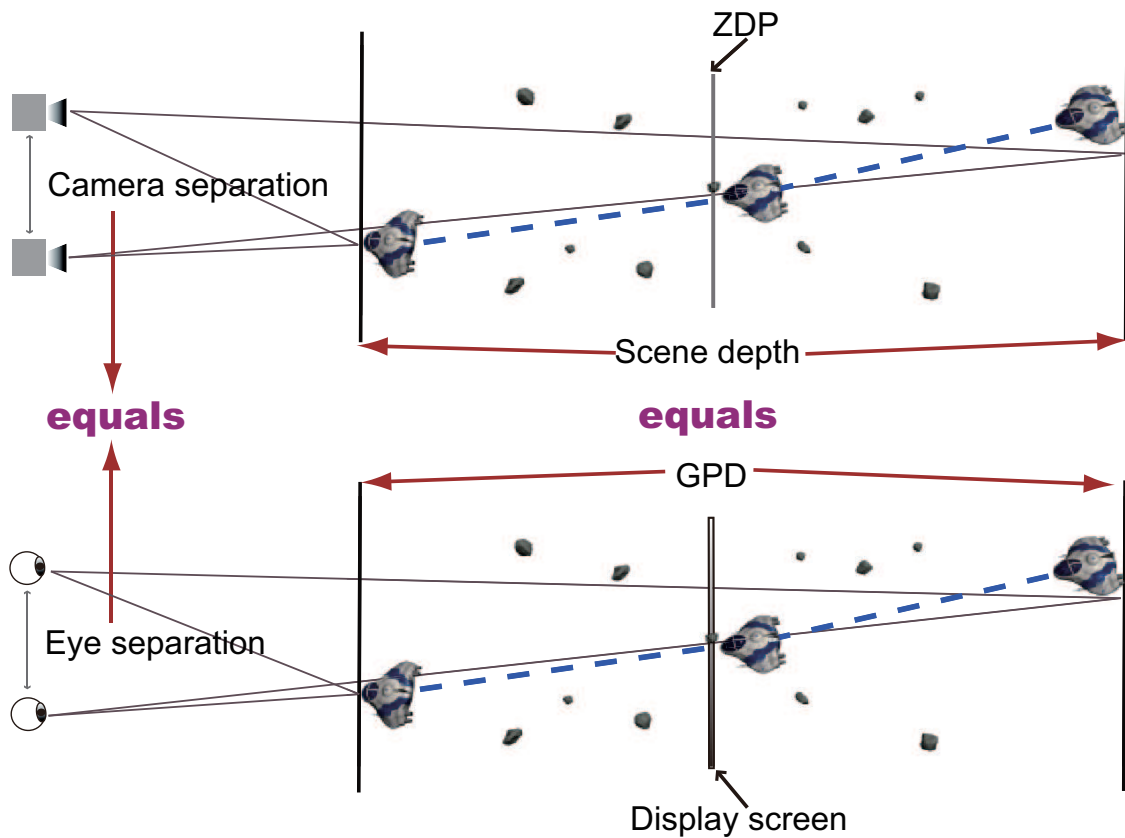


Figure 5.1: Real-Eye Configuration. The camera separation is set equal to the nominal human eye interpupillary distance to simply map the whole scene on top of the display.

This method was included as the comparison baseline for comparing different perceived depth control methods.

5.1.2 Fixed Depth Mapping Approach

We implemented Jones *et al.*'s algorithm [Jones et al., 2001] as the fixed depth mapping method. All figures and equations described in this subsection are from their work.

The position of the viewer, the configuration of the display, and the details of the scene are given in Figure 5.2. There are three unknown variables: Z' , the distance from the cameras to the virtual display which is also known as the Zero Disparity Plane (ZDP); W' , the field width at Z' ; and A , the camera separation. All of them can be calculated from known parameters with a given Field of View (FOV). When using real cameras, the focal length f , and the film with, W_f , must also be specified.

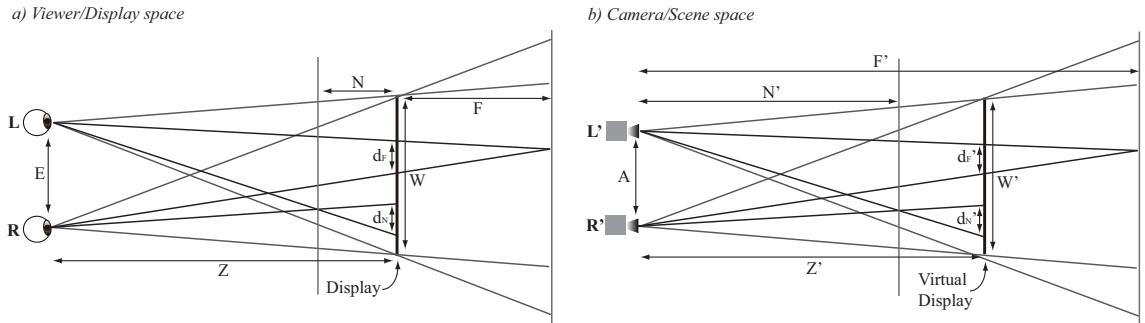


Figure 5.2: Jones *et al.*'s fixed mapping method [Jones et al., 2001]. Notations: a) L and R represent the left and right eyes. E is the interpupillary distance. Z is the viewing distance from the viewer to the display. N is the furthest distance at which objects should appear in front of the screen and F is the furthest distance where objects should appear behind the screen. W is the width of the physical display screen. d_N and d_F are the screen disparities of objects appearing at the distances of N and F. b) L' and R' represent the left and right cameras. A is the camera separation. Z' is the distance from the cameras to the virtual display in the scene. N' is the distance between the closest visible point and the cameras. F' is the distance from the furthest visible point to the cameras. d'_N and d'_F are the world disparities of objects at the distances of N' and F' away from the cameras in the scene.

First we need to have expressions for the screen disparities d_N and d_F :

$$d_N = \frac{NE}{Z - N} \quad \text{and} \quad d_F = \frac{FE}{Z + F} \quad (5.1)$$

and the world disparities, d'_N and d'_F :

$$d'_N = \frac{A(Z' - N')}{N'} \quad \text{and} \quad d'_F = \frac{A(F' - Z')}{F'} \quad (5.2)$$

Equation 5.2 has two unknown variables: Z' and A .

Although the screen disparities and the world disparities are not equal, they should be in same proportion so that the final image is correctly displayed on the screen. We have Z' :

$$\frac{d_N}{d_F} = R = \frac{d'_N}{d'_F} = \frac{(Z' - F')F'}{(F' - Z')N'} \quad (5.3)$$

Equation 5.3 removes A and allows Z' to be calculated as:

$$Z' = \frac{R + 1}{\frac{1}{N'} + \frac{R}{F'}} \quad (5.4)$$

substituting equations 5.1 and 5.3 into equation 5.4, we have:

$$Z' = \frac{\frac{d_N}{d_F} + 1}{\frac{1}{N'} + \frac{\frac{d_N}{d_F}}{F'}} = \frac{\frac{ZN+FN}{ZF-NF} + 1}{\frac{1}{N'} + \frac{ZN+FN}{(ZF-NF)F'}} \quad (5.5)$$

The depth mapping between the scene space and the display space has now been established. Objects between N' and Z' units from the cameras in the scene should be N units in front of the display, and objects between Z' and F' units from the cameras in the scene should be F units behind the display once the correct camera separation is obtained.

As the FOV is given and Z' has been obtained, the width of the field at Z' , W' , can be calculated:

$$W' = 2Z' \tan \frac{\theta}{2} = \frac{Z'W_f}{f} \quad (5.6)$$

Equation 5.5 provides a scaling from the physical display to the virtual display:

$$s = \frac{W'}{W} \quad (5.7)$$

The camera separation, A , can be calculated, given $d'_N = Sd_N$, from:

$$A = \frac{d'_N N'}{Z' - N'} = \frac{Sd_N N'}{Z' - N'} = \frac{\frac{W'}{W} d_N N'}{Z' - N'} \quad (5.8)$$

Replacing W' , d_N , and Z' with equations 5.6, 5.1, and 5.5, we have all the parameters required by the OpenGL viewing frustum now. The frustum can be specified as follows, note that all parameters must be multiplied by $\frac{\text{near clip plane distance}}{Z'}$. Regarding the left camera the viewing frustum should be:

$$l = -(\frac{W'}{2} - \frac{A}{2}) \quad r = \frac{W'}{2} + \frac{A}{2} \quad (5.9)$$

$$t = \frac{H'}{2} \quad b = -\frac{H'}{2} \quad (5.10)$$

and similarly for the right viewing frustum when the viewer sits in front of the center of the display screen.

When capturing 3D images with real cameras, the symmetric viewing frustum is imposed, which requires increasing the FOV and cropping the images to maintain the same field width of the left and right images. A new FOV, focal length, and cropping fraction can be calculated as follows:

$$\theta' = 2 \arctan \frac{W' + A}{2Z'} \quad f' = \frac{W_f}{2 \tan \frac{\theta'}{2}} = \frac{W_f Z'}{W' + A} \quad crop = \frac{A}{W' + A} \quad (5.11)$$

crop refers to the image proportion to be cropped from the left side of the left image and the right side of the right image.

When it is impossible to adjust the focal length by the small amount specified by equation 5.11, the camera separation can be calculated differently. The actual field width captured is:

$$W' + A = 2Z' \tan \frac{\theta}{2} \quad (5.12)$$

which can alter the scale factor to:

$$S = \frac{W' + A}{W + \frac{A}{S}} \quad (5.13)$$

rearranging equation 5.13 we have:

$$\frac{A}{S} = \frac{d_N N'}{Z' - N'} \quad (5.14)$$

substituting equation 5.14 into equation 5.13 and then into equation 5.8, the camera separation A is calculated:

$$A = \frac{2Z' \tan \frac{\theta}{2} d_N N'}{W(Z' - N') + d_N N'} \quad (5.15)$$

Consider equation 5.8 and equation 5.15, Z' is the distance of the cameras from the virtual display (ZDP) in the scene; W is the width of the physical display screen; W' is the field width in the scene which can be calculated by using equation 5.6; θ is the field of view of the camera frustum; N' is distance from the cameras to the closest visible point in the scene; d_N is the crossed screen disparity, which can be derived using equation 5.1 where N , E , and Z are known variables. All of the parameters can be specified subjectively by the viewer.

This method can guarantee any object in the scene appear inside the limits of the 3D Comfort Zone and no excessive stereoscopic depth is perceived. However, the scene can be very large in real-time graphics or 3D cinema despite the fact that sometimes only a small fraction of its volume is actually occupied. Therefore, the perceived depth could be much smaller than the limits of the 3D Comfort Zone when there is a substantial difference between the limits of the scene depth range and the actual occupied volume of scene depth, as demonstrated in Figure 5.3(a).

5.1.3 New Dynamic Depth Mapping Approach

We developed a new dynamic depth mapping method based on Jones *et al*'s fixed depth mapping method. Our new method can automatically adjust the camera separation according to the actually occupied scene depth in real time, so that viewer's perceived depth stays constant on the display always utilises the whole volume of the limits of the 3D Comfort Zone. The sudden and dramatic perceived depth change on the same side of the display is also eliminated by maintaining a constant perceived depth.

Consider Figure 5.3, there is a substantial difference between the maximum scene depth range and the actually occupied volume of scene depth, e.g. the spaceship is at A or B. The perceived depth with the fixed depth mapping method can be much smaller than the predefined limits as illustrated in Figure 5.3(a). With our dynamic mapping approach illustrated in Figure 5.3(b), the perceived depth stays constant on the display and always occupies the whole volume of the 3D Comfort Zone.

The real-time update of camera separation can be achieved by Off-screen rendering using Frame Buffer Object and the Z-buffer value in OpenGL.

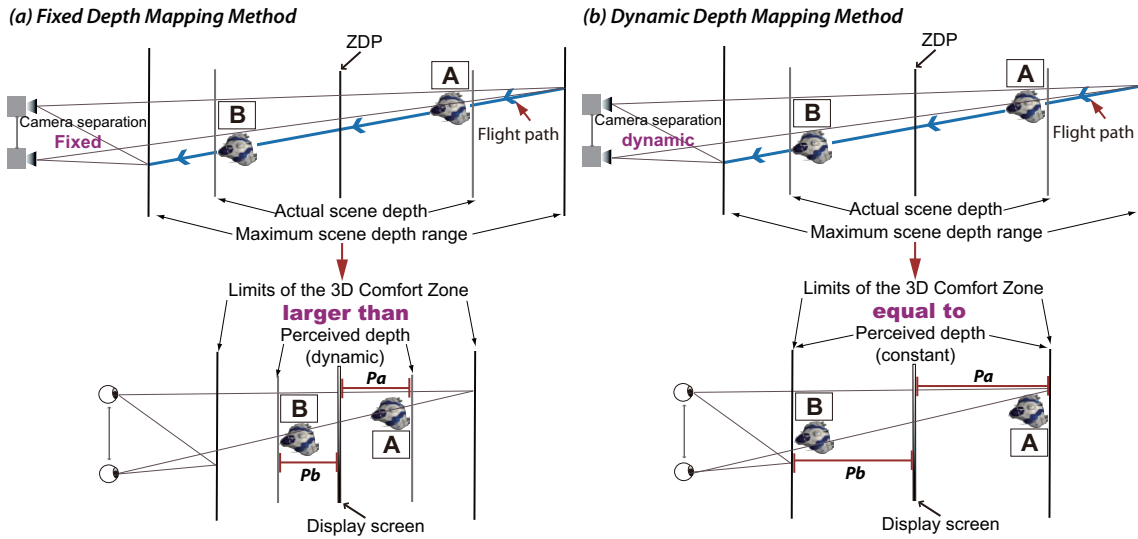


Figure 5.3: Jones *et al*'s fixed mapping method vs. our dynamic mapping method

Frame Buffer Object

The **Frame Buffer Object** (FBO) is an extension to the OpenGL. The FBO is capable of flexible off-screen rendering, such as texture rendering. By drawing images which are normally rendered to the display screen, it can be employed to implement different post-processing effects and image filters. In this method, it was used to achieve the off-screen rendering. The procedures taken to set up the FBO is illustrated in the pseudo-codes below.

Implementation of Z-Buffer

The Z-Buffer also known as depth buffer in OpenGL stores every pixel's depth information frame by frame in the form of a two-dimensional array (horizontal resolution * vertical resolution). This array can be accessed through:

```
glEnable(GL_DEPTH_TEST);
glReadPixels(x, y, width, height, format, type, *pixels);
```

x, y specify the window coordinates of the first pixel that is read from the frame buffer; $width, height$ specify the dimensions of the pixel rectangle; $format$ specifies the format of the pixel data; $type$ specifies the data type of the pixel data and $pixels$

Algorithm 1 Initiating the Frame Buffer Object

```

1: procedure INITFRAMEBUFFER(fbo, depthBuffer)
2:   // Setup our FBO
3:   glGenFramebuffersEXT(1, &fbo);
4:   glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, fbo);

5:   // Create the render buffer for depth
6:   glGenRenderbuffersEXT(1, &depthBuffer);
7:   // Bind depthBuffer to current render buffer
8:   glBindRenderbufferEXT(depthBuffer);
9:   // Ask OpenGL to give storage space of the size of the depth buffer
10:  glRenderbufferStorageEXT(width, height);

11:  // Attach the depth render buffer to the FBO
12:  glFramebufferRenderbufferEXT(depthBuffer);

13:  // Check the status of FBO
14:  GLenum status = glCheckFramebufferStatusEXT();
15:  if status  $\neq$  GL_FRAMEBUFFER_COMPLETE_EXT then
16:    Exit; // Exit if not complete
17:  end if
18: end procedure

```

stores the Z-Buffer value.

In order to convert the Z-Buffer value into the scene depth value, we first need to convert the range of Z-Buffer value from $[0, 1]$ to $[-1, 1]$ by:

$$zValue_{new} = zValue_{old} \times 2 - 1 \quad (5.16)$$

Given (5.16), the scene depth value can be derived from inverting projection matrix on the Z coordinate:

$$SceneDepthValue = \frac{2 \times farZ \times nearZ}{zValue_{new} \times (farZ - nearZ) - (farZ + nearZ)} \quad (5.17)$$

$nearZ, farZ$ are the distances to the near and far clipping planes.

Replacing N' in equation 5.8 or 5.15 with $SceneDepthValue$, the camera separation, A , can be calculated based on the actual occupied scene depth volume for every frame rendered using the off-screen rendering technique in OpenGL. Although this real-time update of camera separation brings extra computational costs, modern graphics system should still be able to render the scene smoothly.

The following pseudo-codes describe how to update scene boundaries with real-time scene depth information from the Z-Buffer.

Algorithm 2 Updating Scene Boundaries with Z-Buffer

```

1: function UPDATE_SCENE_BOUNDARIES(pixelHeight, pixelWidth, farZ, nearZ)
2:   int height = pixelHeight, width = pixelWidth;      // Resolution in pixels
3:   int fz = farZ, nz = nearZ;      // Distances to far and near clipping plane
4:   int numPixels = height * width;
5:   GLfloat *depth = new GLfloat[numPixels];
6:   double minSceneDepth =  $+\infty$ , maxSceneDepth =  $-\infty$ ;

7:   // Read the Z-Buffer value
8:   glReadPixels(0, 0, width, height, GL_DEPTH_COMPONENT, GL_FLOAT, depths);

9:   for i = 0 to i < numPixels do
10:    // Converting Z-Buffer value to scene depth value
11:    double z = depth[i] * 2.0 - 1.0;
12:    double sceneDepth =  $2 * fz * nz / (z * (fz - nz) - (fz + nz))$ ;
13:    // Replacing Max and Min with current scene depth value
14:    if sceneDepth < minSceneDepth then
15:      minSceneDepth = sceneDepth;
16:    else if sceneDepth > maxSceneDepth && sceneDepth < farZ then
17:      maxSceneDepth = sceneDepth;
18:    end if
19:  end for

20:  // Update scene boundaries
21:  sceneBoundaries_near = minSceneDepth;
22:  sceneBoundaries_far = maxSceneDepth;
23: end function

```

Off-screen Rendering using FBO and Z-Buffer

The off-screen rendering is employed to retrieve the real-time scene depth information. The FBO needs to be attached to the current rendering buffer so that the monoscopic image can be rendered to it. The actual scene depth values obtained from the Z-Buffer are then used to replace the predefined scene boundaries. The step-by-step procedure is demonstrated below.

Algorithm 3 Off-screen Rendering

```

1: procedure DRAWOFFSCREEN()
2:   // First we bind the FBO so we can render to it
3:   glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, fbo);

4:   // Save the view port and set it to the size of the buffer
5:   glPushAttrib(GL_VIEWPORT_BIT);
6:   glViewport(0, 0, bufferWidth, bufferHeight);

7:   Draw monoscopic scene...

8:   UpdateSceneBoundaries();

9:   // Restore old view port and set rendering back to default render buffer
10:  glPopAttrib();
11:  glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, 0);
12: end procedure

```

5.1.4 Depth of Field Blur Simulation

As discussed in the preceding chapter, literature suggested that the DoF blur simulation could have advantages in controlling perceived depth in S3D scenes [Blohm et al., 1997, Yano et al., 2004]. We implemented two DoF simulation approaches, one with a deep and fixed DoF centered around the display and another with a shallow and dynamic DoF following the moving object.

Fixed Depth of Field Simulation

The Fixed DOF effect is simulated using multi-pass rendering and accumulation buffer in OpenGL. This method, acting like a real camera lens, has a fixed volume of Depth of Field inside which objects appear in full sharpness. Objects that fall out of the DOF are blurred. The level of blurring is proportional to the distance from the object to the focus plane, which is the ZDP in this work as shown in Figure 5.4. The camera separation used in this approach is also the nominal eye separation, 65 mm.

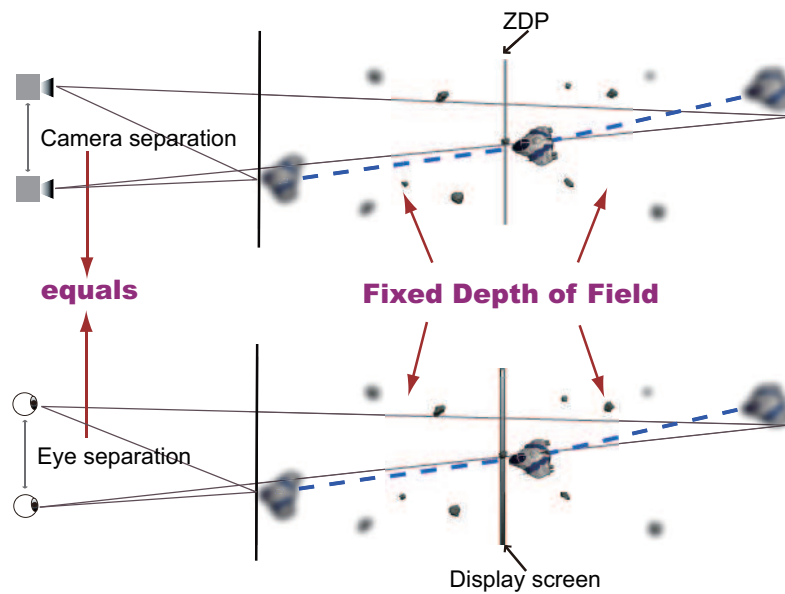


Figure 5.4: The fixed DOF blur method uses the real-eye camera separation and has a fixed volume of Depth of Field inside which objects can be seen in full sharpness.

Dynamic Depth of Field Simulation

The dynamic DoF simulation implemented for this study is similar to the one described by Blohm *et al* [Blohm et al., 1997] in the way that it also has a dynamic DoF following a flying object. We included this method to evaluate the benefits of the dynamic DoF without asking viewers to focus on any specific scene object.

Regarding Figure 5.5, this method too uses multi-pass rendering and accumulation buffer in OpenGL with a real-eye camera separation. It has a dynamic DoF which follows the spaceship. As a result, the spaceship, the only moving object in

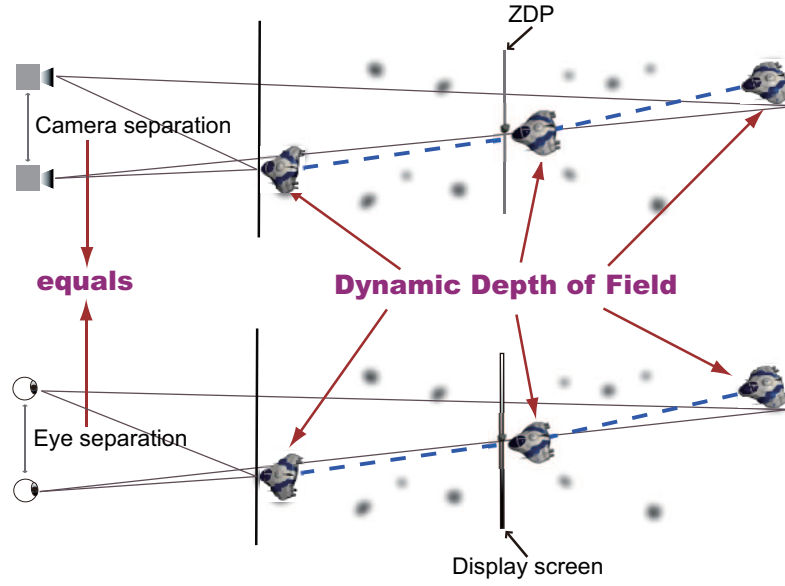


Figure 5.5: The dynamic DOF blur method also uses the real-eye camera separation but has a dynamic Depth of Field which follows the flying spaceship. Objects coming close to the spaceship are also seen in full sharpness.

the scene, always stays inside the DOF. The static objects (asteroids) go in and out of focus depending on their distances to the plane of spaceship.

5.2 Experiment

5.2.1 Protocol

Objective, Method, Comparison Baseline, and Hypothesis

In order to provide a baseline for choosing stereoscopic imaging method to control depth perception in stereoscopic motion pictures, we conducted a subjective human-based experiment to assess the quality of stereoscopic video sequences generated by the five different perceived depth control methods discussed in the preceding section. The experiment followed the Single Stimuli with Multiple Repetition (SSMR) method from the ITU Recommendation BT.500-11 [Union, 2002].

The ITU Recommendation is a widely used methodology for the subjective assessment of the quality of television pictures [Speranza et al., 2006]. It describes methods for measuring the quality of stereoscopic sequences (BT.500). The reason

we chose the SSMR method was that only a single video sequence was presented to the viewer at one time in our experiment.

The comparison baseline for evaluating different methods was the Real-Eye Configuration method.

Our hypothesis was that both depth mapping algorithms and DOF blur effect simulations would have an advantage over the Real-Eye Configuration method in controlling depth perception in 3D motion pictures. We also expected that the video sequence created by our new dynamic depth mapping approach would have a higher quality score than the one generated by the fixed mapping algorithm. We were not certain about viewer's preference between the fixed and dynamic DOF simulations.

Subjects

Seventeen subjects, fifteen male and two female, took part in the experiment. Their ages varied from 20 to 32 with a mean of 24 years. Subjects were not aware of the purpose of the experiment and they were all non-expert, in that their normal work do not concern stereoscopic graphics. All the participants had near-normal vision 20/30 or better tested using a Snellen chart as recommend by the ITU protocol in [Union, 2002], stereo-acuity at 40 sec-arc and passed the colour vision test. Subjects who normally wore glasses or contact lenses were asked to wear them to take the vision tests and during the experiment. These requirements were set to minimize the impact of the difference in human visual systems on the experiment results. All subjects received a nominal payment of five pounds.

Procedure

The experiment consisted of two parts: training session and test session. The experiment began with a training session which demonstrates the range and the type of the scenarios to be assessed. Five video sequences (generated by five different methods) played in the training session were different from those played in the test session, but of comparable sensitivity. The results from the training session were not taken into account in the results analysis.

The procedure of test session is demonstrated in Figure 5.6. There were three

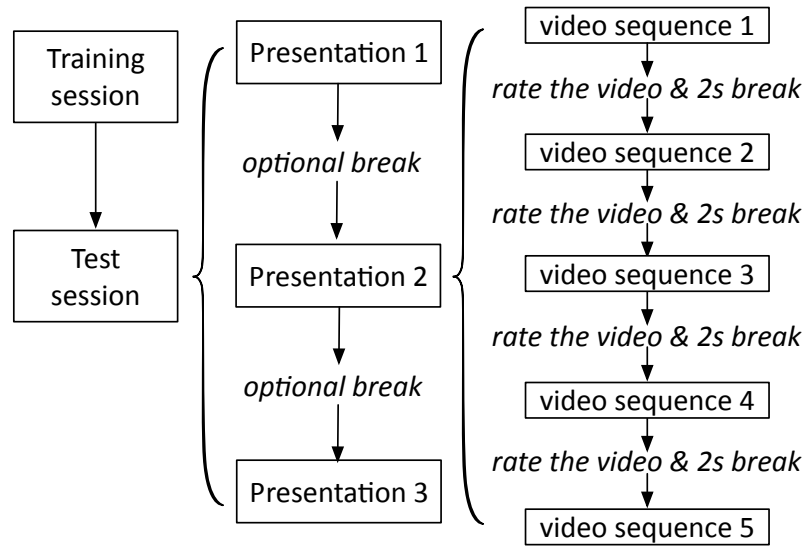


Figure 5.6: Experiment Procedure

presentations as specified by the ITU Recommendation. Each presentation included all five video sequences generated by five methods only once. Each subject watched $5 \times 3 = 15$ video sequences. Each video sequence lasted 20 seconds. The first presentation was used to stabilise the viewer's rating of the video sequences. The data obtained from this presentation were not included in the results of the test. The scores assigned to the video sequences were obtained by taking the mean of the data obtained from the last two presentations.

Each subject watched the video sequences in a different order which satisfied the following limitations specified by the ITU Recommendation: “a given video sequence was not located in the same position in the other presentations; a given video sequence was not immediately located before the same sequence in the other presentations [Union, 2002].”

The video sequences were played on a stereoscopic display. A 2D display was used to show the quality scoring sliders and the subject was asked to record and submit his/her results via this screen. After the submission of the score, there was a two-second blank interval of gray displayed on the 3D display, so that the subject could have time to direct his/her focus back on the 3D screen before the next video sequence was played.

All subjects kept their chins on a chin rest during the whole experiment, as



Figure 5.7: Experiment Apparatus

illustrated in Figure 5.7. The chin rest was 70 cm in front of the 3D display, resulting in a 70 cm viewing distance.

All subjects were given the chance to ask questions before, during and after the experiment and understood they were free to withdraw from the experiment at any time. Subjects were interviewed when they finished. Comments made by subjects were recorded in a text document.

The three vision tests took about 10 minutes. The training session took about 5 minutes and the test session lasted about 15 minutes including optional breaks after each presentation.

Scoring Scale

Figure 5.8 illustrates the scoring scale shown on the 2D screen. The video sequences were rated on a sliding scale of Excellent, Good, Fair, Poor, and Bad. These terms categorised the five different levels and they were the same as those normally used in the ITU-R recommendation. The terms were associated with the value intervals of 100 to 80, 80 to 60, 60 to 40, 40 to 20, and 20 to 0, respectively, providing a continuous rating system. Subjects were asked to score the quality of each stereoscopic video sequence by moving the slide bar to the desired position along the scale. The vertical scale displayed on the 2D screen was ten centimetres long and divided into

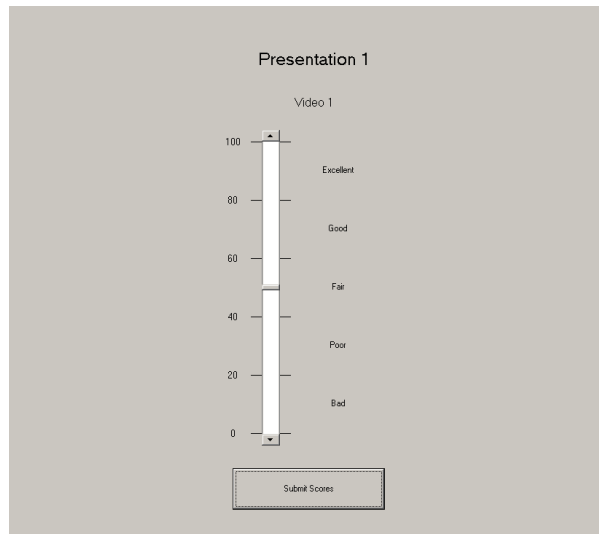


Figure 5.8: Scoring scale

five equal lengths. Results were recorded once subjects clicked the “Submit Scores” button.

Apparatus and Viewing Conditions

The experiment was run by a Dell Precision PWS670 computer with Intel Xeon CPU of 3.00GHz 2.99 GHz, 4.00 GB RAM and NVIDIA Quadro FX5600 graphics card. A 24-inch True 3Di stereoscopic display with a 1920 × 1200 resolution was used to play the experimental test video sequences. This 3D display required viewers to wear polarised glasses to fuse the left and right images. The scoring scale was shown on a 21-inch HITACHI CRT with a resolution of 1280 × 1024. We chose a CRT monitor for displaying the scoring scale so that viewers did not need to take on/off glasses when switching between 3D and 2D displays.

The two displays run independently. However, both of them used the graphics card from NVIDIA Quadro FX family and were driven by NVIDIA ForceWare Release 80. The whole experiment was conducted in a quiet dimly lit room.

Stimulus

The five different stimulus tested in the experiment are shown in Figure 5.9, 5.10, 5.11, 5.12 and 5.13. The test scene was an animation generated by Computer Graph-

ics (CG). They all had the same frame composition. The scene was composed of a flying spaceship and eleven still asteroids (one of them is occluded by the spaceship in the figures). The spaceship was flying back and forth through those asteroids along a curvilinear path resembling a figure of 8. Its velocity varied smoothly along the flight trajectory. The spaceship would slow down as it turned around to avoid any undesired visual artifacts that could be caused by sharp and sudden turns. The scene was designed to have both still and moving objects to test if the viewer would spontaneously focus on the moving object. In our experiment, subjects were not required to keep their focuses on any particular scene object.

Figure 5.12 demonstrates the Fixed DOF stimuli. The spaceship is heavily blurred as it is quite far away from the focus plane on which the central asteroid is located. Figure 5.13 illustrates the Dynamic DOF stimuli. We can see that the asteroid at bottom-right corner is also inside the DOF and seen in full sharpness due to the close distance between itself and the plane of the spaceship. Note that the 3D models of spaceship and asteroids are existing models supplied by the Visualisation Lab at Durham University.



Figure 5.9: Real-Eye Stimuli

3D Studio Max was used to model the CG animated scene, edit the flight path of the spaceship and generate its coordinates. Our software read in the exported scene objects and flight path coordinates and rendered the scene frame by frame in OpenGL. Each test stimuli had 500 frames in total with a frame rate of 25 fps. Viewer's maximum perceived depth of Real-Eye Configuration, Fixed DOF



Figure 5.10: Fixed Mapping Stimuli



Figure 5.11: Dynamic Mapping Stimuli



Figure 5.12: Fixed DOF Stimuli

and Dynamic DOF stimulus was 200 mm in front of the display screen and 240 mm behind the display screen, same as the scene depth maximum. The viewer's



Figure 5.13: Dynamic DOF Stimuli

maximum perceived depth of Fixed and Dynamic Mapping approaches was 50 mm and 60 mm on each side of the display.

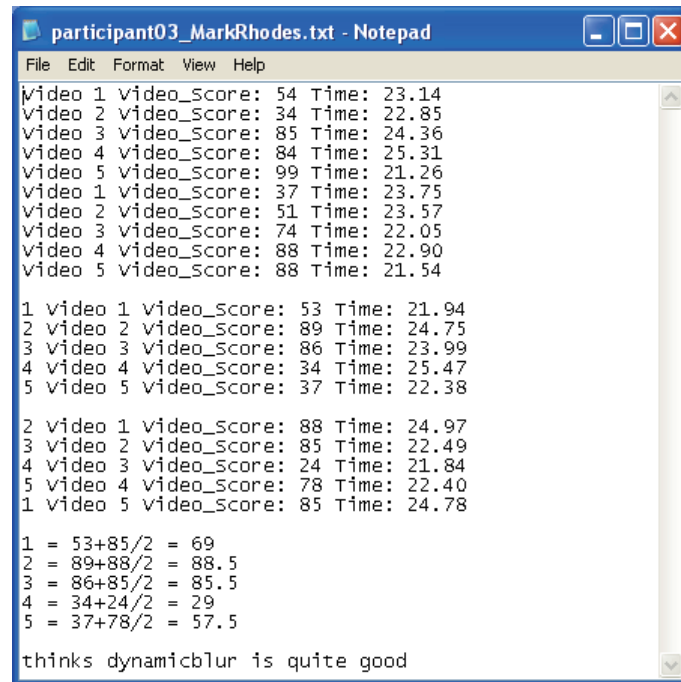
5.2.2 Results and Analysis

A result sheet is demonstrated in Figure 5.14. The first ten results are obtained from the training session and are not taken into account in the result analysis. 1, 2, 3, 4, and 5 are designated to the video sequences generated by the real-eye configuration, fixed depth mapping, dynamic depth mapping, fixed DOF blur and dynamic DOF method respectively, followed by the video sequence number in each test session, the score assigned by the viewer, and the time spent by the viewer to rate the video sequence. The mean score of each video sequence is automatically calculated. Viewers' comments are also included.

Discussion of the Results

Figure 5.15 demonstrates the scores of each method tested in the experiment (Real-Eye: the real-eye method, FixedMap: the fixed depth mapping method, DynamicMap: the dynamic depth mapping method, FixexBlur: the fixed DoF method, DynamicBlur: the dynamic DoF method).

Figure 5.16 is the box plot of the results from all seventeen subjects. Consider each method, the dashed line (first line from bottom) is the sample minimum which is the smallest number among all seventeen scores; the bold line (first line from top)



```

participant03_MarkRhodes.txt - Notepad
File Edit Format View Help
video 1 video_Score: 54 Time: 23.14
video 2 video_Score: 34 Time: 22.85
video 3 video_Score: 85 Time: 24.36
video 4 video_Score: 84 Time: 25.31
video 5 video_Score: 99 Time: 21.26
video 1 video_Score: 37 Time: 23.75
video 2 video_Score: 51 Time: 23.57
video 3 video_Score: 74 Time: 22.05
video 4 video_Score: 88 Time: 22.90
video 5 video_Score: 88 Time: 21.54

1 video 1 video_Score: 53 Time: 21.94
2 video 2 video_Score: 89 Time: 24.75
3 video 3 video_Score: 86 Time: 23.99
4 video 4 video_Score: 34 Time: 25.47
5 video 5 video_Score: 37 Time: 22.38

2 video 1 video_Score: 88 Time: 24.97
3 video 2 video_Score: 85 Time: 22.49
4 video 3 video_Score: 24 Time: 21.84
5 video 4 video_Score: 78 Time: 22.40
1 video 5 video_Score: 85 Time: 24.78

1 = 53+85/2 = 69
2 = 89+88/2 = 88.5
3 = 86+85/2 = 85.5
4 = 34+24/2 = 29
5 = 37+78/2 = 57.5

thinks dynamicblur is quite good

```

Figure 5.14: An example of a result sheet

is the sample maximum and the largest score assigned to this method; the bottom of the box is the 25th percentile (lower quartile, Q1) of the results and the top of the box is the 75th percentile (upper quartile, Q3); the line inside the box is the 50th percentile (the median, Q2).

By observing Figure 5.15 and the box plot, we can make the following observations: the real-eye method often performs fairly well, with one outlier who does not like it; the fixed mapping method performs fairly similarly to the real eye method, with similar range of scores and variances; the dynamic mapping method seems to be the best and the most consistent one as it constantly performs well with the highest mean and lowest variance; the fixed blur method constantly performs poorly, with the lowest mean and low variance; the dynamic blur method has a very large variance, which indicates very divergent options. Overall, the dynamic mapping method seems to outperform the rest of the methods both in terms of consistency (lowest variance) and rating (highest mean). The fixed mapping method and the real eye method also perform fairly well. The DoF blur methods are the worst with the dynamic blur clearly outperforms the fixed blur, but is much less consistent.

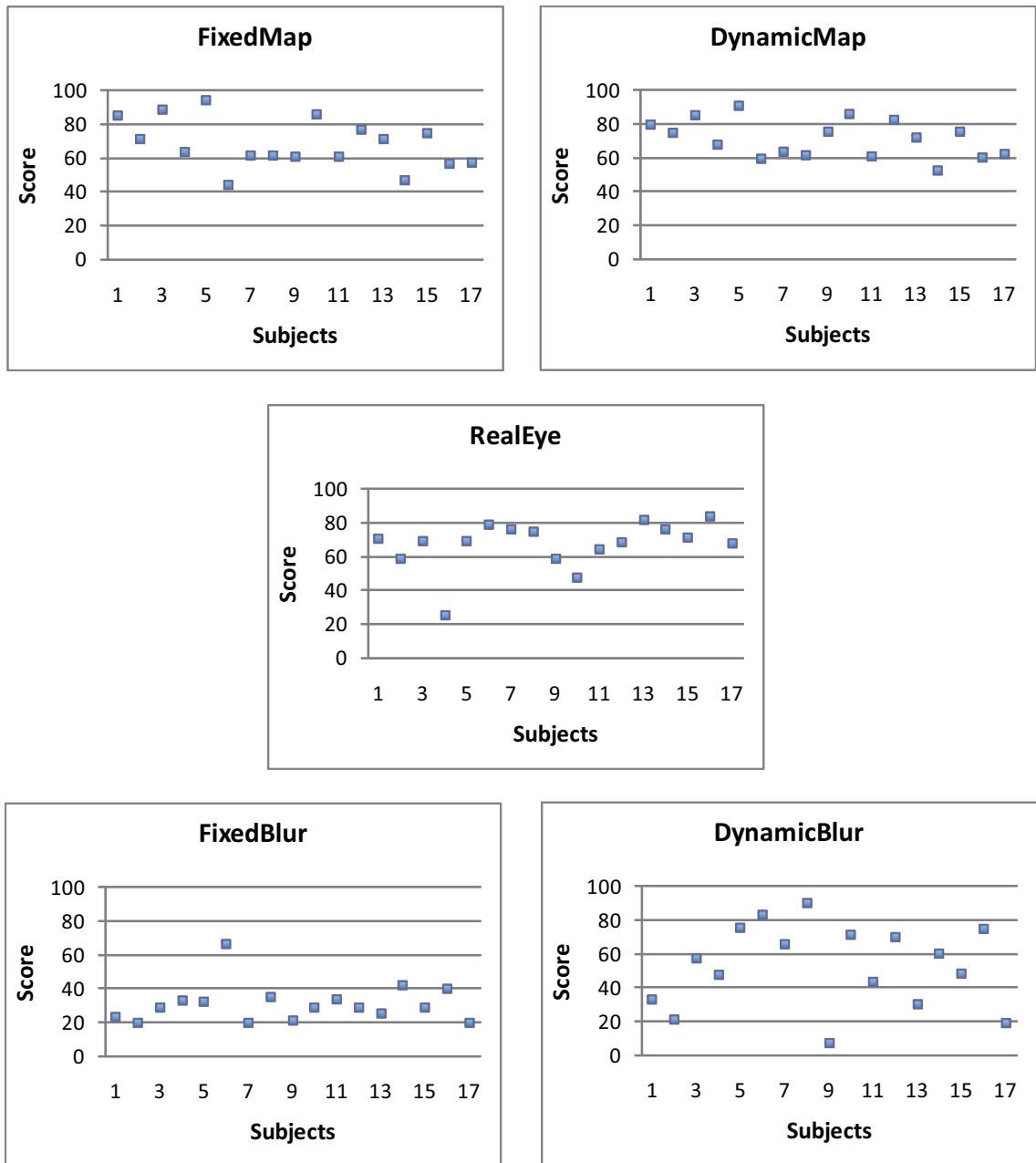


Figure 5.15: Results of each method tested in the experiment

The statistical analysis described below confirmed these observations.

Note that the symbol * represents the outlier. We tracked down the outliers in the results. The outlier for the real-eye method came from Participant 4. In the interview after the experiment, he stated that he really did not like a large perceived depth whilst most of the subjects could cope well with the large perceived depth created by the real-eye method. The outlier for the fixed blur method came from

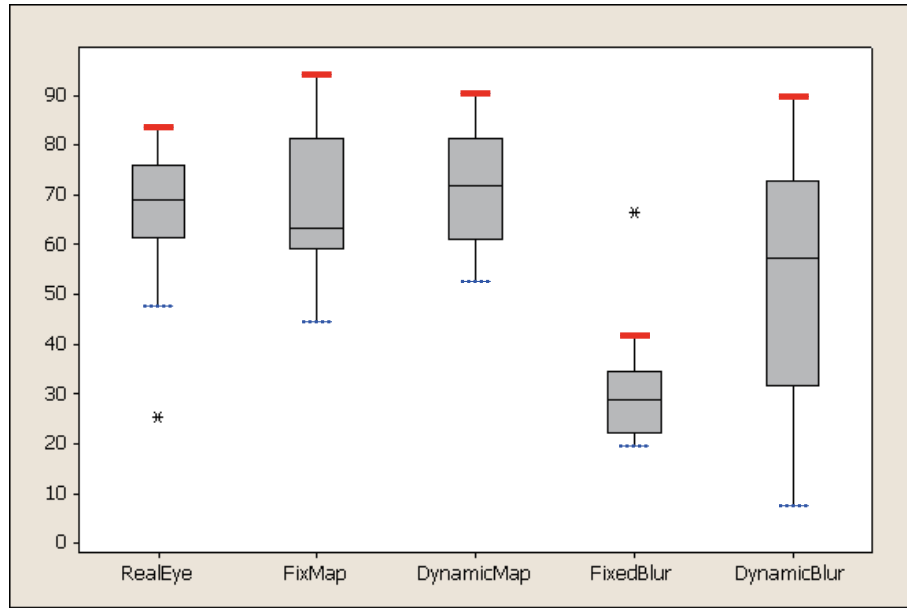


Figure 5.16: Box Plot Results

Participant 6. He stated that he really liked a large perceived depth and the blur effect. This could explain why he preferred the fixed blur method over both depth mapping methods whilst most of the subjects really disliked the fixed blur method as they did not like watching the spaceship blurred.

Statistical Analysis of the Results

The mean score and standard deviation for each method is shown in Table 5.1. Only the mean (μ_{fb}) of Fixed DOF method fell below 50 which corresponded with the term “Fair” in ITU’s grading scale. Viewers seemed to be satisfied with the quality of the video sequences generated by other four methods. The results from one-way Single Factor ANOVA indicated that the differences between different methods were statistically significant, F ratio = 19.117 > F critical = 2.486. A Paired T-Test was then performed on every possible interaction.

The paired t-test is used to compare the values of means from two related samples. In our experiment, the subjects watched the same video sequences, therefore, the scores of one method by all the subjects can be considered as i.i.d (independent and identically distributed) variables. On the other hand, the scores of all methods

Table 5.1: Means and Standard Deviations

Method	Mean, μ	St. Dev.
Real-Eye	$\mu_{re} = 67.088$	13.989
FixedMap	$\mu_{fm} = 68.441$	14.414
DynamicMap	$\mu_{dm} = 71.206$	11.214
FixedBlur	$\mu_{fb} = 31$	11.413
DynamicBlur	$\mu_{db} = 52.882$	24.223

by the same subject may be related. Therefore, we can use paired t-test to compare the means of two methods.

We were interested in testing whether one method was worse than another method (i.e., the mean of the method is smaller than the mean of the other method). Therefore, we used one-sided hypothesis test with the null hypothesis being $H_0 : \mu_1 = \mu_2$ and the alternative hypothesis being $H_1 : \mu_1 < \mu_2$. If we can reject the null hypothesis H_0 , we conclude Method 1 is likely to be worse than Method 2.

The results of the T-Test comparisons concerning the comparison baseline: Real-Eye Configuration is listed in Table 5.2.

Table 5.2: T-Test results concerning the Real-Eye method

Hypotheses	p value	Conclusion
$H_0 : \mu_{re} = \mu_{fm}$ vs. $H_1 : \mu_{re} < \mu_{fm}$	0.401	Fail to reject H_0
$H_0 : \mu_{re} = \mu_{dm}$ vs. $H_1 : \mu_{re} < \mu_{dm}$	0.206	Fail to reject H_0
$H_0 : \mu_{re} = \mu_{fb}$ vs. $H_1 : \mu_{re} < \mu_{fb}$	1.000	Fail to reject H_0
$H_0 : \mu_{re} = \mu_{db}$ vs. $H_1 : \mu_{re} < \mu_{db}$	0.985	Fail to reject H_0

As shown in Table 5.1, the mean scores of Real-Eye, Fixed Mapping and Dynamic Mapping methods (μ_{re} , μ_{fm} , μ_{dm}) were quite similar. Paired T-Tests were carried out and only differences between Fixed Mapping and Dynamic Mapping approaches were statistically significant, $H_0 : \mu_{fm} = \mu_{dm}$ vs. $H_1 : \mu_{fm} < \mu_{dm}$, p value = 0.028. Neither Dynamic mapping nor Fixed mapping method was able to provide

the expected advantage over the Real-Eye approach in controlling depth perception in stereoscopic cinematography, as shown in Table 5.2. Only two subjects out of seventeen reported that they did not like the stereoscopic videos generated by Real-Eye method as they provided too much perceived depth. All other subjects had no problem fusing the stereo pairs with perceived depth that was four times as large as the limits of the 3D Comfort Zone suggested by Jones *et al.* These results suggested that the perceived depth range tested: 20 cm in front of and 24 cm behind the display were inside the 3D Comfort Zone on our tested display.

Table 5.3: T-Test results between mapping methods and DOF simulations

Hypotheses	p value	Conclusion
$H_0 : \mu_{fm} = \mu_{dm}$ vs. $H_1 : \mu_{fm} < \mu_{dm}$	0.028	Reject H_0
$H_0 : \mu_{fb} = \mu_{db}$ vs. $H_1 : \mu_{fb} < \mu_{db}$	0.000	Reject H_0
$H_0 : \mu_{fb} = \mu_{fm}$ vs. $H_1 : \mu_{fb} < \mu_{fm}$	0.000	Reject H_0
$H_0 : \mu_{fb} = \mu_{dm}$ vs. $H_1 : \mu_{fb} < \mu_{dm}$	0.000	Reject H_0
$H_0 : \mu_{db} = \mu_{fm}$ vs. $H_1 : \mu_{db} < \mu_{fm}$	0.019	Reject H_0
$H_0 : \mu_{db} = \mu_{dm}$ vs. $H_1 : \mu_{db} < \mu_{dm}$	0.000	Reject H_0

Regarding Table 5.3, the results from a Paired T-Test comparison between Fixed and Dynamic DOF methods revealed that there was a 100% probability that the difference between Dynamic DOF simulation and Fixed DOF simulation was statistically significant, $H_0 : \mu_{fb} = \mu_{db}$ vs. $H_1 : \mu_{fb} < \mu_{db}$, p value = 0.000. Paired T-Tests were also performed between the two depth mapping methods and the two DOF simulations. The results, listed in Table 5.3, showed that the differences in the means between the following pairs: dynamic mapping and fixed mapping, fixed mapping and dynamic blur, dynamic blur and fixed blur, were positive, and these differences were statistically significant. These results indicated that dynamic mapping method was the best method among the depth mapping and DOF simulation methods, while fixed mapping method being the second highest, dynamic blur method being the third, and fixed blur method being the worst. To summarise, the depth mapping methods outperformed the DOF simulations, and the dynamic (depth mapping and

DoF) methods outperformed the corresponding fixed (depth mapping and DoF) methods.

5.3 Conclusion

We performed a subjective human-based experiment to evaluate different methods for controlling the perceived depth in stereoscopic cinematography.

Our new approach of dynamic mapping of depth from the scene space to display space scored the highest mean among all the tested approaches. Statistics confirmed that it was able to provide a significant effect over the fixed mapping algorithm and DOF simulations in controlling the perceived depth in stereoscopic cinematography. It is our assumption that its lack of statistical significance over the real-eye configuration was due to the relatively conservative assignment of the scene depth tested.

We also learned that, in contrast to the conclusions drawn by previous studies, the DOF blur simulation does not improve the perceived depth quality in 3D cinematography. However, there were indications in our results suggesting that viewers could regard the Dynamic DOF simulation as a good imitation of natural visual experience when there are both dynamic and static objects in the scene.

Detailed analysis of the practical benefits of the dynamic depth mapping method and the DOF blur techniques are presented in Chapter 7.

Chapter 6

Identifying the 3D Comfort Zone on a Desktop 3D Display

In this chapter, we propose a novel method that employs the *Random Dot Stereogram* (RDS) technique to identify the range of the 3D Comfort Zone on a given desktop stereoscopic display. The details of the experiment conducted to evaluate the method and the statistical analysis of the experiment results are also provided.

6.1 Random Dot Stereogram

The Random Dot Stereogram technique was invented by Dr. Bela Julesz [Julesz, 1971]. A Random Dot Stereogram, shown in Figure 6.1, is a pair of images of random dots, which produces the sensation of depth with objects appearing to be in front of or behind the actual image when viewed with the aid of a stereoscopic viewing device, such as a stereoscope, a pair of stereoscopic glasses, etc. The basic process used to generate a RDS image is: create a background image of suitable size as the left image, fill it with random dots and duplicate the image as the right image; select a region in the left image and laterally shift the region by a certain amount; then the RDS image is complete. If the selected region in the left image is shifted to the right of that region on the right image, the object appears to be in front of the background image; otherwise, the object falls behind the actual image. The depth at which the object appears behind or in front of the background image depends on

the distance between the two identical regions on the left and right images.

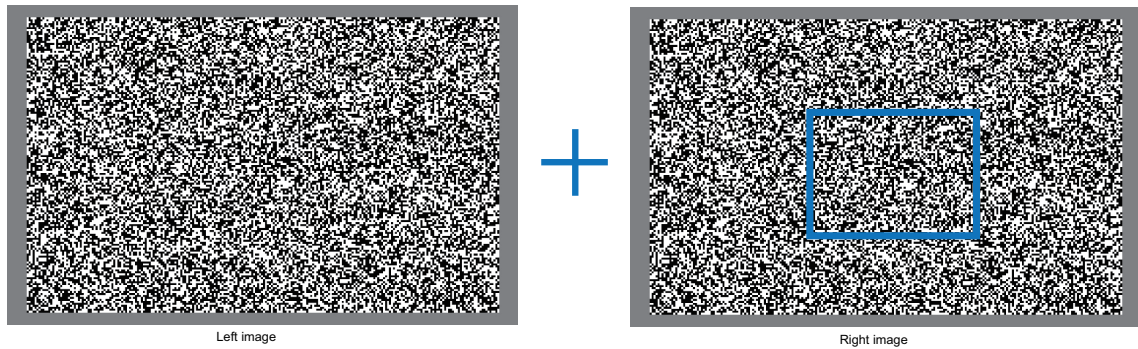


Figure 6.1: Random Dot Stereogram

6.1.1 Random Dot Autostereogram

Another type of RDS images which are normally seen in Magic Eye books are called Random Dot Autostereogram. A Random Dot Autostereogram, illustrated in Figure 6.2, is a single-image random dots stereogram that is similar to a normal RDS image except it is viewed without a stereoscope device. When viewing a normal RDS image with a stereoscopic viewing device, a pair of 2D images of the same scene are presented to the left and right eyes from slightly different angles which allows the brain to reconstruct the original scene based on binocular disparity. When viewing a Random Dot Autostereogram, repeating 2D patterns are presented to the brain from both eyes. However, the brain is unable to correctly match them. By forcing the eyes to converge at a point in front of/behind the actual image, the brain can be tricked to match two adjacent identical patterns into a virtual object using wrong parallax angles, thus perceive the virtual object at a different depth plane from that of the actual image. The distance at which this plane falls behind/in front of the actual image depends on the horizontal distance between the identical patterns. In Figure 6.2, a shark should emerge from the background if one can achieve the required eye vergence.

All the Random Dot Stereograms used in this study are normal RDS images which do not require viewers to force the eye vergence. With the aid of 3D display and glasses, viewers with proper stereo vision should be able to recognise the hidden

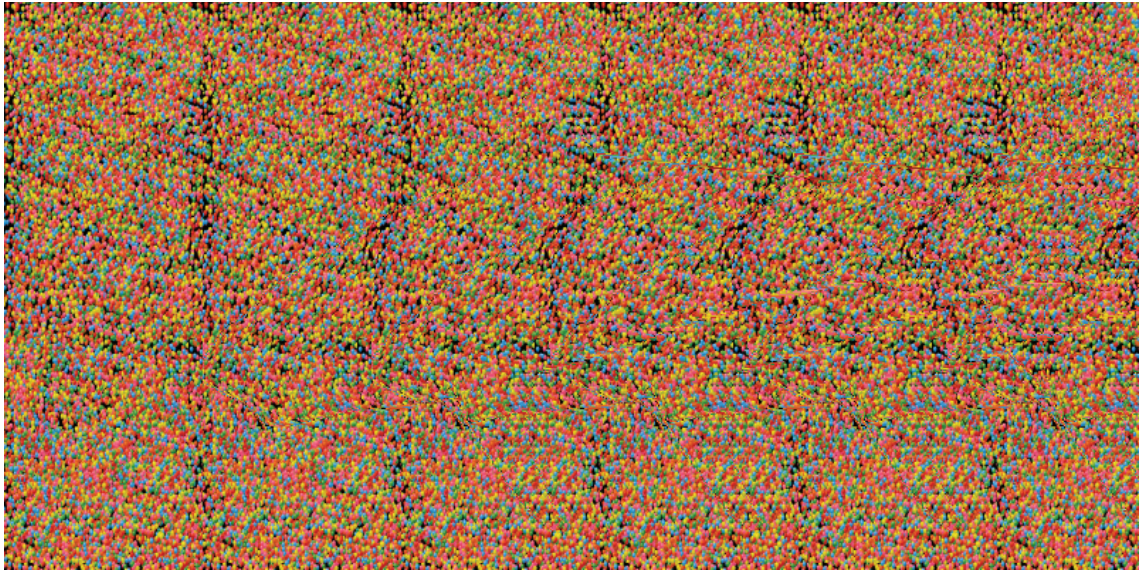


Figure 6.2: Random Dot Autostereogram [aut, 2011]

3D scene easily and straightforwardly. The concept of Random Dot Autostereogram is only introduced as background information. We refer to only normal Random Dot Stereograms when using the term RDS in this study. The reason we chose the RDS technique to define the practical viewing range around a 3D display is that RDS images impose viewers to achieve correct fusion of them in order to identify the 3D scene hidden in the random dots. Hence, it is easy to draw conclusions on the perceived depth limits around a 3D display based on when viewers lose fusion and fail to distinguish the 3D scene from the background image.

6.1.2 Generating RDS Images with the Stencil Buffer in OpenGL

Creating random dots images in PhotoShop and shifting the desired pattern manually could be the most straightforward way to generate a stereoscopic RDS image. However, in order to make our method self-contained, we programmed a method that automatically draws and presents stereoscopic RDS images using the Frame Buffer and Stencil Buffer in the OpenGL. There are three main functions in the method: (1) attaching a texture to the Frame Buffer. (2) drawing a RDS image onto the texture. (3) using Stencil Buffer to mask out a portion of the background

texture image. A detailed description of those functions is provided in the following sections.

Attaching a texture to the Frame Buffer

The introduction of the Frame Buffer has been given in Chapter 5 and we took the same steps to set up the Frame Buffer Object (FBO) and Depth Buffer as we did in Chapter 5. However, instead of rendering to the Frame Buffer, we had to render to a texture which requires the creation of a texture and attaching it to the FBO. In order to successfully create a texture in OpenGL, one must start with naming it using `glGenTextures` and binding it to a texturing target by `glBindTexture`. Then, the specifications and parameters of the texture is given by `glTexImage2D` and `glTexParameterf`. Lastly, the texture is attached to the FBO using `glFramebufferTexture2DEXT`.

Drawing a RDS image onto a texture

Drawing RDS images in OpenGL is rather easy and straightforward as demonstrated in the pseudo-codes below. Given the size (in pixel) of a dot, the number of dots in one row and column can be derived by `ImageWidth/DotSize` and `ImageHeight/DotSize`. The colour of each dot is randomised. Before the rendering takes place, the pixel storage mode needs to be set with `glPixelStorei` and the raster position for pixel operations is specified with `glRasterPos2i`. Finally, `glDrawPixels` is called to write the pixels to the Frame Buffer and subsequently onto the texture.

Stencil Buffer

The Stencil Buffer is an extra buffer in OpenGL in addition to the depth buffer (Z-Buffer) and the pixel (colour) buffer. The Stencil Buffer is typically used to limit the rendering area (stenciling) by masking out a portion of the image plane in OpenGL. The Stencil Buffer is initiated by `glEnable` with the argument `GL_STENCIL_TEST`, and controlled with the function `glStencilFunc` which sets function and reference

Algorithm 4 Drawing RDS Image

```

1: function DRAWRDSIMAGE(imageHeight,imageWidth,RDSImage,dotSize)
2:   int dotColour;
3:   int i, j, k, m, n, x;                                     // variables for the loops
4:   for i = 0 to i < imageHeight/dotSize do
5:     k = dotSize * i;
6:     for j = 0 to j < imageWidth/dotSize do
7:       n = dotSize * j;
8:       int ranNum = rand()%100 + 1;                          // ranNum ∈ [1, 100]
9:       if ranNum ≤ 50 then
10:        dotColour = 255;                                     // Dot Colour is white
11:      else
12:        dotColour = 0;                                       // Dot Colour is black
13:      end if
14:      for m = 0 to m < dotSize do
15:        for x = 0 to x < dotSize do                      // Assign colour to each dot
16:          RDSImage[m + k][n + x][0] = dotColour;
17:          RDSImage[m + k][n + x][1] = dotColour;
18:          RDSImage[m + k][n + x][2] = dotColour;
19:        end for
20:      end for
21:    end for
22:  end for
23: end function

```

value for the stencil testing. Actions taken based on the outcome of the stencil test are specified with `glStencilOp`.

Algorithm 5 Rendering Stereoscopic RDS image with Stencil Buffer

```

1: procedure RENDERSTEREORDSIMAGE
2:   if rendering left frame then
3:     Draw RDS image;
4:   else
5:     Draw RDS image;
6:     glTranslated(1.0,0.0,0.0);           // Horizontally shift the viewing matrix

7:     // Enable Stencil Buffer and set it to 1 where we draw the pattern
8:     glEnable(GL_STENCIL_TEST);
9:     glStencilFunc(GL_ALWAYS,1,1);
10:    glStencilOp(GL_KEEP, GL_KEEP, GL_REPLACE);

11:    Draw the pattern;

12:    // Draw only where Stencil Buffer is 1 and keep it unchanged
13:    glStencilFunc(GL_EQUAL,1,1);
14:    glStencilOp(GL_KEEP, GL_KEEP, GL_KEEP);

15:    Redraw the RDS image;
16:   end if
17: end procedure

```

As shown in the pseudo-codes above, suppose we choose to shift the pattern in the right image. All needed to be done when rendering the left image is to draw the random dot image. As for rendering the right image, firstly we need to draw an identical random dot image as the one for the left image so that the two images can have the same textured background; secondly the Stencil Buffer is enabled so the region where the pattern is drawn can be masked; after shifting the view and drawing the pattern, we need to redraw the same random dot image again with the

Stencil Buffer remaining the same so the newly drawn pattern is not overshadowed by the background. Then a RDS image is complete.

6.2 Experiment

In order to identify the practical limits of the 3D Comfort Zone on a given desktop 3D display, we conducted a human-based experiment where RDS images were presented to viewers at different perceived depth planes both in front of and behind the display screen, and subjects were asked to identify the 3D shapes hidden in the RDS images. We drew our conclusion on the practical limits of the 3D Comfort Zone according to the accuracy and time taken on the identifications.

6.2.1 Protocol

Hypothesis

Based on the previous studies on the limits of the 3D Comfort Zone (discussed in Section 4.1), we made the following seven hypotheses:

H1: For the RDS image perceived *in front of* the display:

Subjects should be able to recognise the hidden shapes *correctly* when the image is presented on a perceived depth plane inside the practical limits of the 3D Comfort Zone. They should make more errors as the perceived depth plane of the image increase. At extremely large perceived depth planes, subjects might not be able to recognise the shapes.

H2: Similar to H1, for the RDS image perceived *behind* the display:

Subjects should be able to recognise the hidden shapes *correctly* when the image is presented on a perceived depth plane inside the practical limits of the 3D Comfort Zone. They should make more errors as the perceived depth plane of the image increase. At extremely large perceived depth planes, subjects might not be able to recognise the shapes.

H3: For the RDS image perceived *in front of* the display:

Subjects should be able to recognise the hidden shapes *quickly* when the image is presented on a perceived depth plane inside the practical limits of the 3D Comfort Zone. They should take more time as the perceived depth plane of the image increase. However, at extremely large perceived depth planes, since they may be purely guessing, the time subjects take may decrease instead.

H4: Similar to H3, for the RDS image perceived *behind* the display:

Subjects should be able to recognise the hidden shapes *quickly* when the image is presented on a perceived depth plane inside the practical limits of the 3D Comfort Zone. They should take more time as the perceived depth plane of the image increase. However, at extremely large perceived depth planes, since they may be purely guessing, the time subjects take may decrease instead.

H5: For the RDS image perceived *in front of* the display:

The subjects should have similar speed and accuracy in recognising the two hidden shapes (*square or triangle*).

H6: Similar to H5, for the RDS image perceived *behind* the display:

The subjects should have similar speed and accuracy in recognising the two hidden shapes (*square or triangle*).

H7: The limit of the 3D Comfort Zone *in front of the display* should be smaller than the limit *behind* the display.

Subjects

Twenty-four subjects, fifteen male and nine female, took part in the experiment. Their ages varied from 23 to 34 with a mean of 25 years. Subjects were not aware of the purpose of the experiment and they were all non-expert, in that their normal work do not concern stereoscopic graphics. All subjects met the minimum criteria of acuity of 20:30 vision and stereo-acuity at 40 sec-arc. Subjects who normally wore glasses or contact lenses were asked to wear them to take the vision tests and during the experiment (the purpose of setting these requirements was explained in the subsection: Subjects in Section 5.2.1). As the experiment did not require viewing

colourful contents, colour vision test was not performed on the subjects. All subjects volunteered to participate and there was no payment for the participation of the experiment.

Apparatus and Viewing Conditions

The apparatus and viewing conditions of this experiment were the same as those of the experiment described in the Chapter 5 except that the 2D CRT display was not employed.

Experiment Procedure

The experiment consisted of two parts, pre-test session and test session. The pre-test session was designed to check if subjects who had passed the stereo vision test could actually identify hidden shapes from RDS images. All five RDS images in the pre-test session were presented on the display so the hidden shapes were just in front of the screen. Only subjects who could identify all the shapes in the pre-test session were allowed to proceed with the experiment.

In the test session each subject was presented with a stereoscopic RDS image, representing either a square or triangle, on the stereoscopic display. Once the subject had identified the shape, he/she was instructed to press the spacebar on the keyboard. A menu (illustrated in Figure 6.3), consisting of two labelled buttons “square” and “triangle”, would appear in the middle of the screen for the subject to choose the shape he/she identified from the RDS image. After the subject had recorded his/her choice by clicking one of the labelled buttons, a blank grey screen would be shown for three seconds, relaxing the subject’s eyes before the next image was presented. The time spent by a subject to identify a shape was also recorded. There was no mechanism for playing backwards.

During the test session, the background RDS image was placed on totally 36 different perceived depth planes with the hidden 3D shape floating above the actual RDS image. There were six perceived depth planes in front of the display, spreading from 10 cm above the display to 60 cm above the display with the interval of 10 cm; and thirty of them behind the display, spreading from 10 cm behind the display to

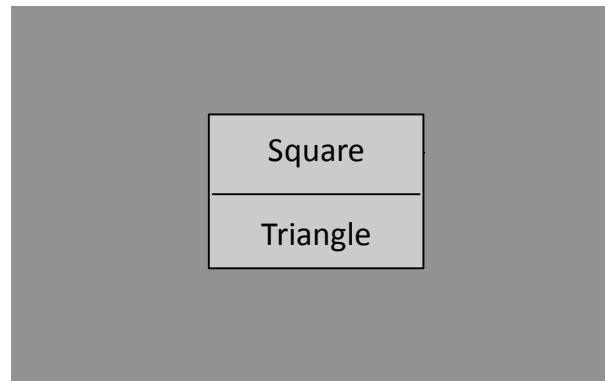


Figure 6.3: Menu for selecting the shape

300 cm behind the display also with the interval of 10 cm. The RDS images and the hidden shapes were displayed in the same size on all tested perceived depth planes.

Each perceived depth plane was viewed by the subject twice, one time for the square and another for the triangle. Each subject saw $36 \times 2 = 72$ RDS images. The order of perceived depth planes being presented was completely random for each subject. The order of the square and triangle appearing on each perceived depth plane for each subject was also randomised. It was designed this way to eliminate the possibility that one of those two shapes is easier to identify than the other. The first five images in the test session were for training purpose and the results were not taken into account in the result analysis.

All subjects kept their chins on a chin rest during the whole experiment, as illustrated in Figure 6.4. The chin rest was 75 cm in front of the 3D display, resulting in a 75 cm viewing distance. Subjects were instructed not to spend too much time on one image. If they could not recognise the shape, they were requested to take a guess and still choose a shape from the menu nonetheless. This was designed to make the statistical analysis doable without bringing any bias to the results.

All subjects were given the chance to ask questions before, during and after the trial and understood they were free to withdraw from the experiment at any time. Subjects were fully debriefed when they finished. The two vision tests and the pre-test session took about 5 minutes and the test session lasted about 10 minutes.



Figure 6.4: Experiment Apparatus

6.2.2 Stimulus

For each tested stereoscopic RDS image, the size of the random dots drawn was 4×4 pixels. The size of the background RDS image was 960×600 pixels. The rest of the screen was covered in a gray colour. The five shapes: circle, hexagon, pentagon, diamond, and star included in the pre-test session are demonstrated in Figure 6.5. The square and triangle included in the test session are demonstrated in Figure 6.6.

Note that the width and height of the RDS image were chosen to be half of those of the display screen to allow easy image perception, and a wide range of screen disparity values being tested without causing partial image perception. The tested perceived depth range was 60 cm in front of the display and 300 cm behind the display. With the given image size and viewing distance, the perceived depth in front of the display could not go any further without causing partial perception of the RDS image; the perceived depth behind the display was limited at 300 cm to avoid the possibility of eye divergence: 300 cm perceived depth behind the display creates 5.1 cm screen disparity, which could be equal to the lower bound of the adult eye separation range according to [Dodgson, 2004]. The perceived depth range was controlled by Jones et al's algorithm [Jones et al., 2001].

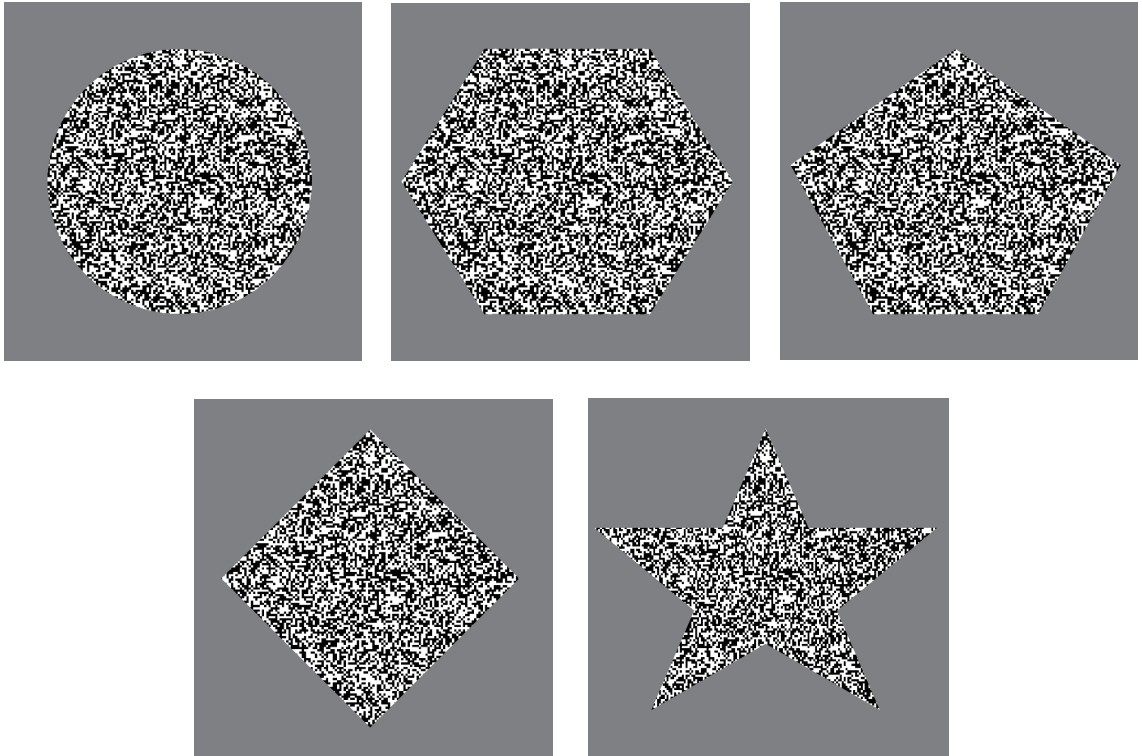


Figure 6.5: Stimulus in the pre-test session

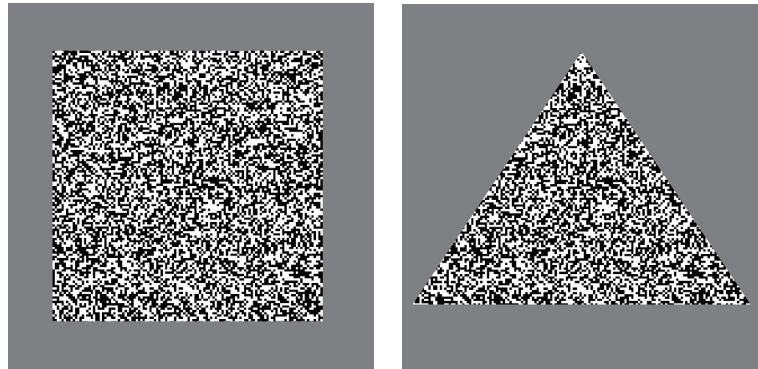


Figure 6.6: Stimulus in the test session

6.3 Results and Analysis

Our software was able to automatically record the correct answer, the choice of viewers', and calculate if the answer was correct (1 being correct and 0 being wrong); In addition the perceived depth (GPD), screen disparity, and the time spent on each image were also recorded, as demonstrated in Figure 6.7. And finally the data were

sorted by an increasing order of the magnitude of screen disparity for in front of the screen and behind the screen respectively.

No	Ans	R/w	GPD	SD	Time
2	2	1	10	1	0.797
1	1	1	10	1	1.062
2	2	1	20	2.364	1.172
1	1	1	20	2.364	1.593
2	2	1	30	4.333	2.296
1	1	1	30	4.333	1.125
1	1	1	40	7.429	2.094
2	2	1	40	7.429	9.531
1	1	1	50	13	10.641
2	2	1	50	13	7.516
1	1	1	60	26	4.734
2	1	0	60	26	4.532

No	Ans	R/w	GPD	SD	Time
1	1	1	10	0.765	0.703
2	2	1	10	0.765	3.39
1	1	1	20	1.368	0.86
2	2	1	20	1.368	0.828
1	1	1	30	1.857	0.593
2	2	1	30	1.857	0.657
1	1	1	40	2.261	0.969
2	2	1	40	2.261	0.625
1	1	1	50	2.6	0.734
2	2	1	50	2.6	0.5
2	2	1	60	2.889	0.625
1	1	1	60	2.889	0.859
1	1	1	70	3.138	0.515
2	2	1	70	3.138	0.796
1	1	1	80	3.355	1.359
2	2	1	80	3.355	0.89
2	2	1	90	3.545	0.625
1	1	1	90	3.545	0.594
1	1	1	100	3.714	0.75
2	2	1	100	3.714	0.485

Figure 6.7: Results

6.3.1 Statistical Analysis Method

On a given depth plane, we generated two pictures, one had a triangle on it, and the other one had a square on it. We asked 24 subjects to identify each shape and give an answer (triangle or square). If an answer is correct, we let $X_i = 1$, otherwise, $X_i = 0$. As a result, we have $n = 24 \times 2 = 48$ data in total. Let the probability with which a person's answer is wrong be p . We can see that X_i 's are Bernoulli distributed, i.e.,

$$X_i = \begin{cases} 0 & w.p \ p \\ 1 & w.p \ 1 - p \end{cases}$$

Let

$$X = \sum_{i=1}^n X_i,$$

then X is binomially distributed, i.e., $X \sim \text{Binomial}(n, p)$.

By de Moivre-Laplace theorem [Papoulis and Pillai, 2002], a binomial distribution is approximately a normal distribution with mean np and standard deviation

$\sqrt{np(1-p)}$, if n is large and certain conditions are satisfied (for example, when p is not too close to 0 or 1). Since $n = 48$ in our experiment, we can also use normal distribution to approximate X .

In our experiment, if a subject is always able to identify the right shapes, $p = 0$. If the subject finds the shape hard to identify, p is some value between 0 and 0.5. If the subject is identifying shapes by pure guess, $p = 0.5$. This is the reason that we asked the subjects to make a guess when they could not identify the shape. If we asked them to claim that they “could not identify the shape” directly, the distribution of the answer would no longer be Bernoulli distributed.

We did “Hypothesis Test” to test which is the case. Our null hypothesis is:

$$H_0 : p = 0.$$

We chose two-proportion analysis as our test method, and 0.05 as the significant level. If the normal approximation test (or Fisher’s (binomial) exact test) reports a p-value that is larger than the significant level, we conclude that the data are consistent with the null hypothesis that $p = 0$, otherwise, we reject the null hypothesis.

Because the normal approximation is valid, we can draw the same conclusion from the confidence interval of the approximated distribution. If 0 falls in the 95% confidence interval, one can conclude that the data are consistent with the null hypothesis that $p = 0$, and otherwise, we conclude that the data are conflicting with the null hypothesis.

6.3.2 Depth plane analysis based on mistakes

In Front of the Display

Figure 6.8 demonstrates the number of mistakes made by viewers on each depth plane in front of the display. We can observe that the total number of mistakes increases as perceived depth plane increases. When the perceived depth plane is only 10 (which is the smallest depth plane that we tested), the number of errors is zero.

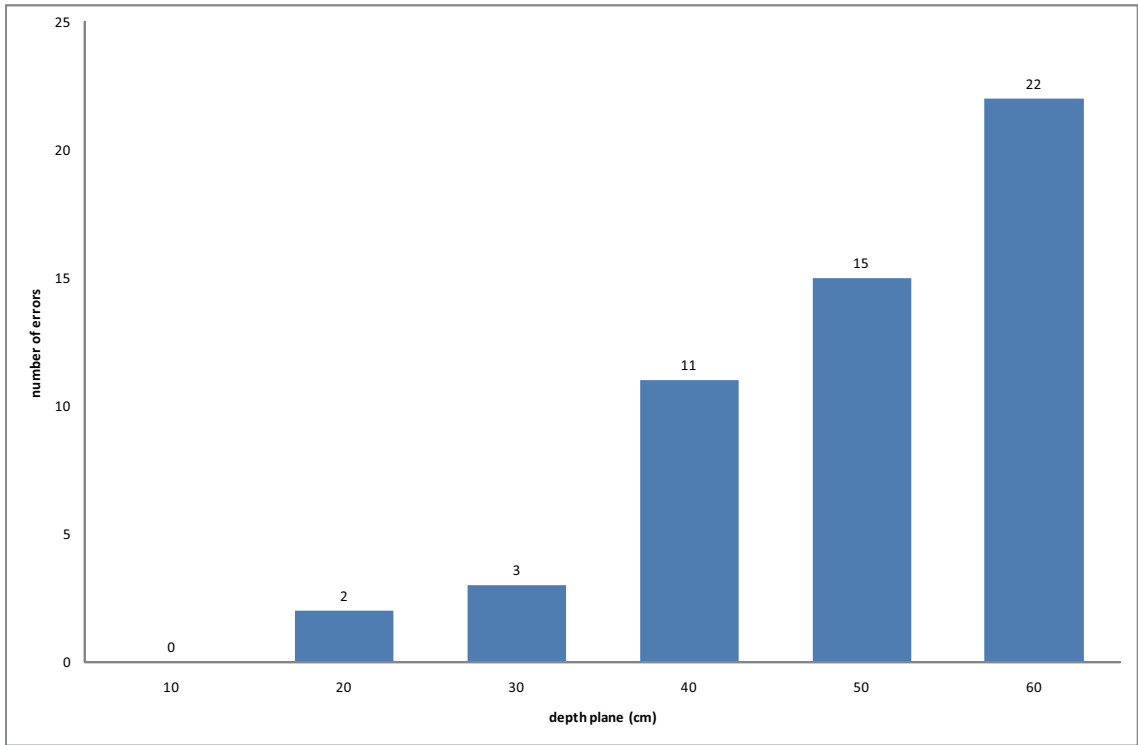


Figure 6.8: Mistakes made on the depth planes in front of the display

As discussed above, we run two proportions (comparison) analysis. The two binomial samples we have are: 1. n_1 trials, x correct choices, probability $p_1 = x/n_1$; 2. n_2 trials, y successes, probability $p_2 = y/n_2$. Our null hypothesis $H_0 : p_1 = p_2$; $H_1 : p_1 \neq p_2$. For each depth plane we have $n_1 = 48$ (24 subjects), $x = 48$ correct choices, $p_1 = 1$. Let $n_2 = 48$, y = the number of correct choices made on each depth plane.

Table 6.1 lists the statistical analysis based on the results. We can observe that our null hypothesis ($p = 0$) cannot be rejected when the perceived depth planes are 10 cm, 20 cm and 30 cm. This implies that the subjects are likely to be able to identify the correct shapes when the perceived depth plane is not above 30 cm. The table also shows that our null hypothesis ($p = 0$) is rejected when the perceived depth planes are 40 cm, 50 cm and 60 cm. This implies that some subjects are unable to identify the correct shapes when the perceived depth plane is above 30 cm. In fact, when the perceived depth plane is 60 cm, the lower bound of the 95% confidence interval (CI) is much greater than 0, and the CI includes 0.5, this implies

Table 6.1: Successes made on the depth planes in front of the screen

Depth	Successes	95% CI	p-value	Significant?
10	48	(*,*)	1.000	No
20	46	(−0.015, 0.098)	0.495	No
30	45	(−0.006, 0.131)	0.242	No
40	37	(0.110, 0.348)	0.001	Yes
50	33	(0.181, 0.443)	0.000	Yes
60	26	(0.317, 0.599)	0.000	Yes

that the subjects were likely to be unable to identify the shapes at all, and their answers were pure guesses.

These results agree with hypothesis H1 and we can conclude that the limit for the 3D Comfort Zone in front of the display should be 30 cm.

Behind the Display

Figure 6.9 demonstrates the number of mistakes made by the viewers on each depth plane behind the display. We can observe that the total number of mistakes is 3 (out of 48) at most, and there is no obvious trend on it as perceived depth plane increases. It implies that the subjects are likely to be able to identify the correct shapes in the whole range of the perceived depth planes that we tested (from 10 cm to 300 cm).

There was no need to run the two proportions analysis for depth planes behind the screen again as there were only four different possibilities: $y = 0, 1, 2, 3$. And even when $y = 3$, there was still no statistical significance as shown in Table 6.1 (Depth = 30 cm). These results indicate the differences are not statistically significant.

These results agree with hypothesis H2. However, we should note that a wider range of the perceived depth planes should be tested in order to check whether the number of mistakes increases with the perceived depth plane increase.

We can conclude that the limit for the 3D Comfort Zone behind the display should be larger than 300 cm. Since we found the limit for the 3D Comfort Zone

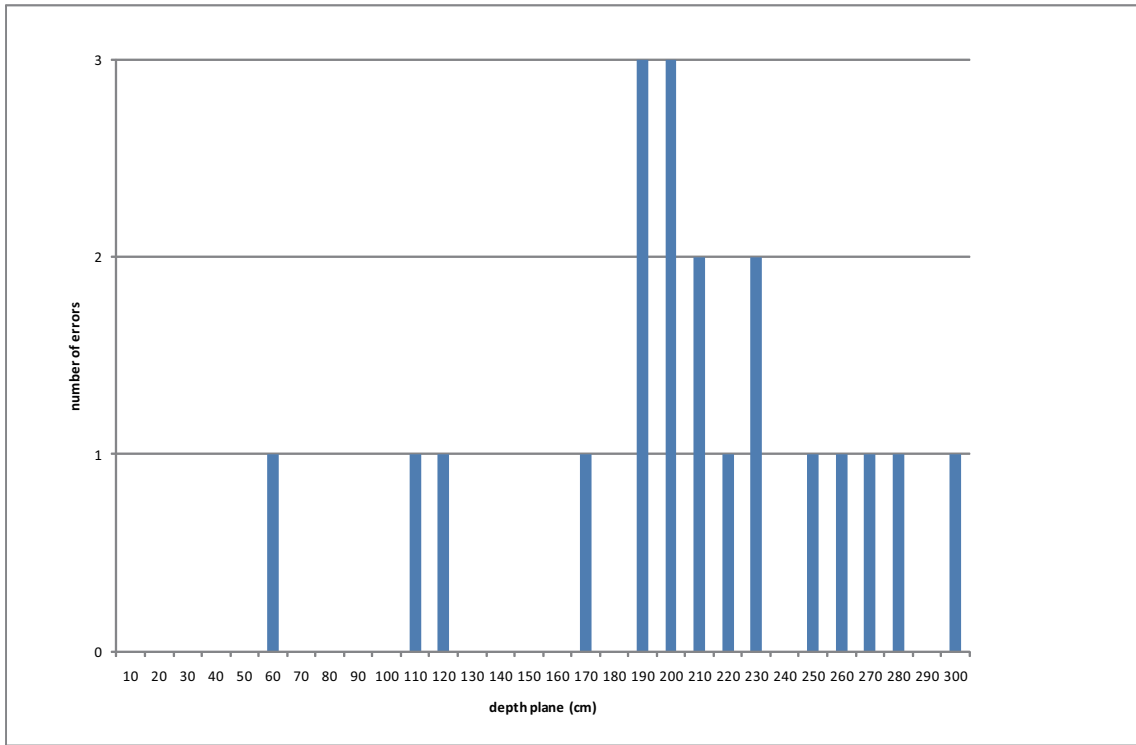


Figure 6.9: Mistakes made on the depth planes behind the display

in front of the display is 30 cm, the results are consistent with hypothesis H7. We expect to define a conservative limit for the perceived depth behind the display based on the time spent on different depth planes.

6.3.3 Depth Plane Analysis based on time

In Front of the Display

Figure 6.10 demonstrates the average time spent on each depth plane in front of the display. We can observe that the time spent on identifying the shapes increase with the perceived depth plane increase when the perceived depth plane goes from 10 cm to 50 cm, however, the time decreases a little when the perceived depth plane goes from 50 cm to 60 cm. As we discussed before, the subjects were likely to be purely guessing the results at depth plane 60 cm, and therefore, they used shorter time than at depth plane 50 cm, at which they might be able to identify the shape with difficulty. At the lowest depth plane 10 cm, the average time spent is only 1.6

seconds.

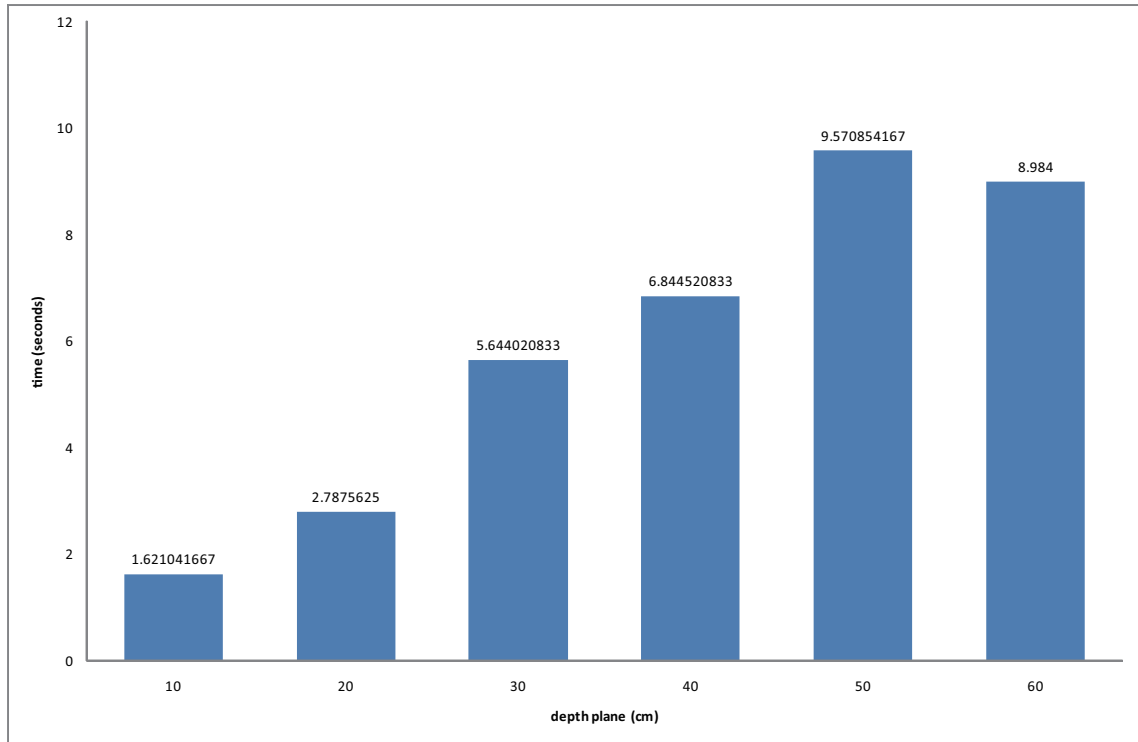


Figure 6.10: Time spent on the depth planes in front of the display

We run paired t-test between the average time of two adjacent depth planes. The results are listed in Table 6.2. It shows that the increases in the time spent are statistically significant within pairs (10 cm, 20 cm), (20 cm, 30 cm), and (40 cm, 50 cm), and not within pairs (30 cm, 40 cm) and (50 cm, 60 cm).

Table 6.2: Time spent on the depth planes in front of the screen

Pair	95% CI	p-value	Significant?
(10, 20)	(−2.259, −0.074)	0.037	Yes
(20, 30)	(−4.810, −0.903)	0.006	Yes
(30, 40)	(−3.80, 1.40)	0.349	No
(40, 50)	(−5.31, −0.14)	0.040	Yes
(50, 60)	(−1.77, 2.95)	0.612	No

The reason we did not do t-test between every possible interaction was we had

already successful defined the depth limit in front of the display through correct choices analysis. The results from the time spent on each depth plane was supplementary only and should only be considered when there was not enough information from the correct choice analysis.

These results agree with hypothesis H3.

Behind the Display

Figure 6.11 demonstrates the average time spent on each depth plane behind the display. We can observe that there is an increasing trend on the time spent on identifying the shapes when the perceived depth plane increases. The time spent on all the perceived depth planes tested ranged from 1.25 seconds (at the smallest depth plane 10 cm) to 2.75 seconds. This implies that the subjects are always fairly quick in identifying the shapes.

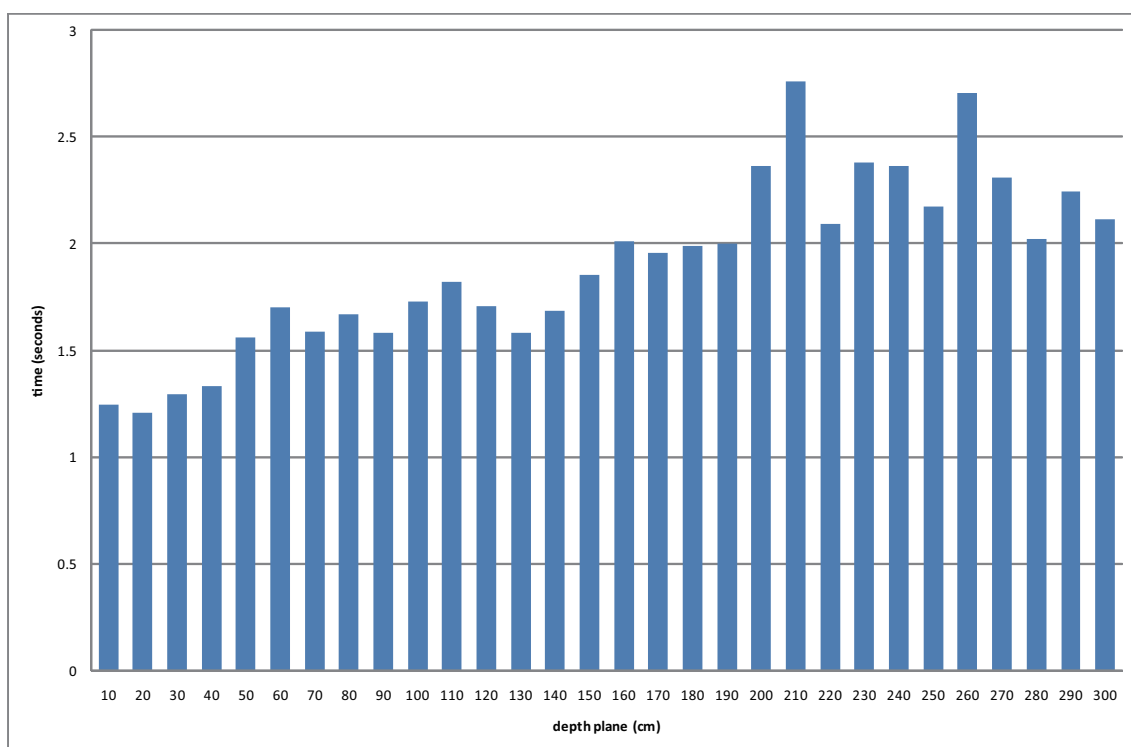


Figure 6.11: Time spent on the depth planes behind the display

We run paired t-test between the average spent on depth plane of 10 cm behind the screen and the average time spent on other depth planes. The average spent on

depth plane of 10 cm is set to be the H_0 . Statistical significance started showing from the depth plane of 50 cm behind the screen as highlighted in Table 6.3. It shows that the increases in the time spent are statistically significant only within pairs (10 cm, 50 cm) and (10 cm, 60 cm), and not within pairs (10 cm, 20 cm), (10 cm, 30 cm) and (10 cm, 40 cm). This indicates that identifying the shapes at perceived depth planes above 50 cm is significantly more difficult than that at perceived depth plane 10 cm.

These results agree with hypothesis H4.

Table 6.3: Time spent on the depth planes behind the screen

Pair	95% CI	p-value	Significant?
(10, 20)	(−0.147, 0.214)	0.703	No
(10, 30)	(−0.227, 0.129)	0.577	No
(10, 40)	(−0.285, 0.104)	0.348	No
(10, 50)	(−0.605, −0.025)	0.035	Yes
(10, 60)	(−0.325, −0.021)	0.015	Yes

In real-time stereoscopic applications such as 3D computer games, subjects are required to perceive the depth quickly and have immediate response. Hence we believe it may be necessary to limit the perceived depth behind the display. Based on our results, we conclude that the conservative perceived depth limit behind the display should be 40 cm.

6.3.4 Comparison between Square and Triangle based on mistakes

In Front of the Display

Figure 6.12 demonstrates the number of wrong choices made for squares and triangles in front of the display. We can observe that the differences in the number of mistakes between the two shapes are very small across all the depth planes. There

are no difference at depth plane 10 cm, 20 cm and 60 cm. And the differences are only ± 1 at depth plane 30 cm, 40 cm and 50 cm.

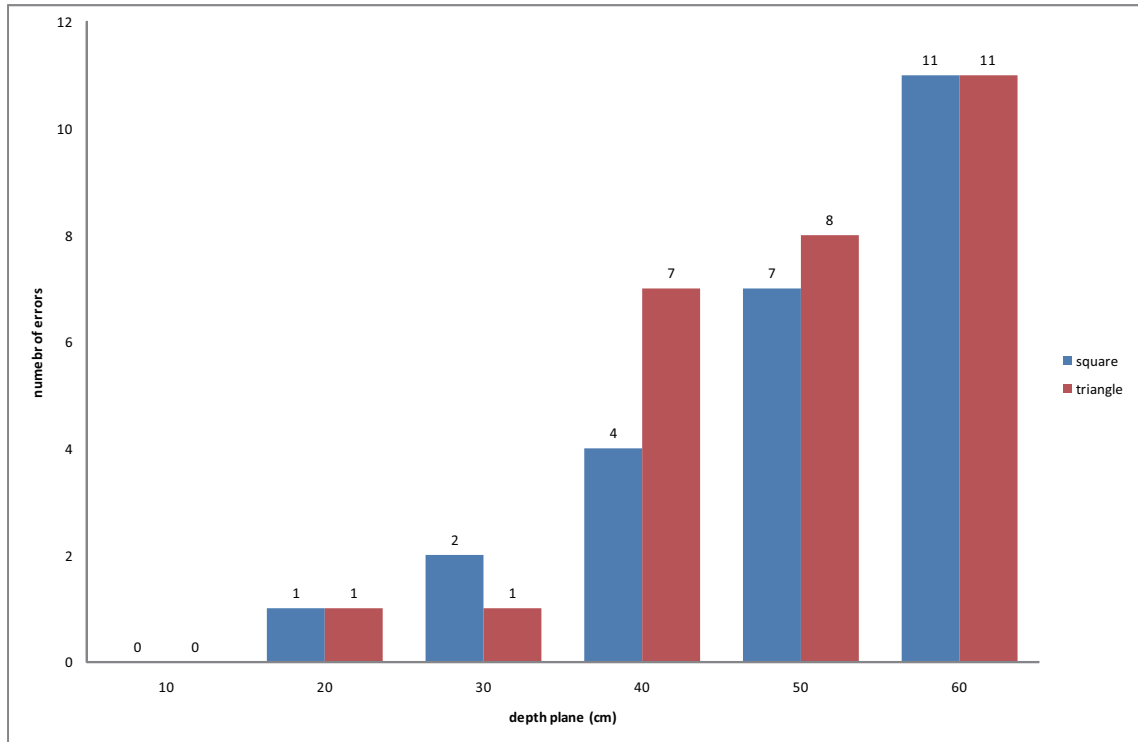


Figure 6.12: Different mistakes made in identifying squares and triangles on the depth planes in front of the display.

From Table 6.4, we can observe that these differences are not statistically significant. The results indicate that the shape of the object does not have an effect on the accuracy. These results agree with hypothesis H5.

Table 6.4: Mistakes made for square and triangle in front of the screen

Depth	Num. of Errors(s t)	95% CI	p-value	Significant?
30	(2 1)	(−0.095, 0.178)	1.000	No
40	(4 7)	(−0.360, 0.110)	0.494	No
50	(7 8)	(−0.304, 0.220)	1.000	No

Behind the Display

Figure 6.13 demonstrates the number of wrong choices made for squares and triangles behind the display. We can observe that the differences in the number of mistakes between the two shapes are very small (0, 1 or 2) across all the depth planes. The results indicate that the shape of the object does not have an effect on our result. These results agree with hypothesis H6.

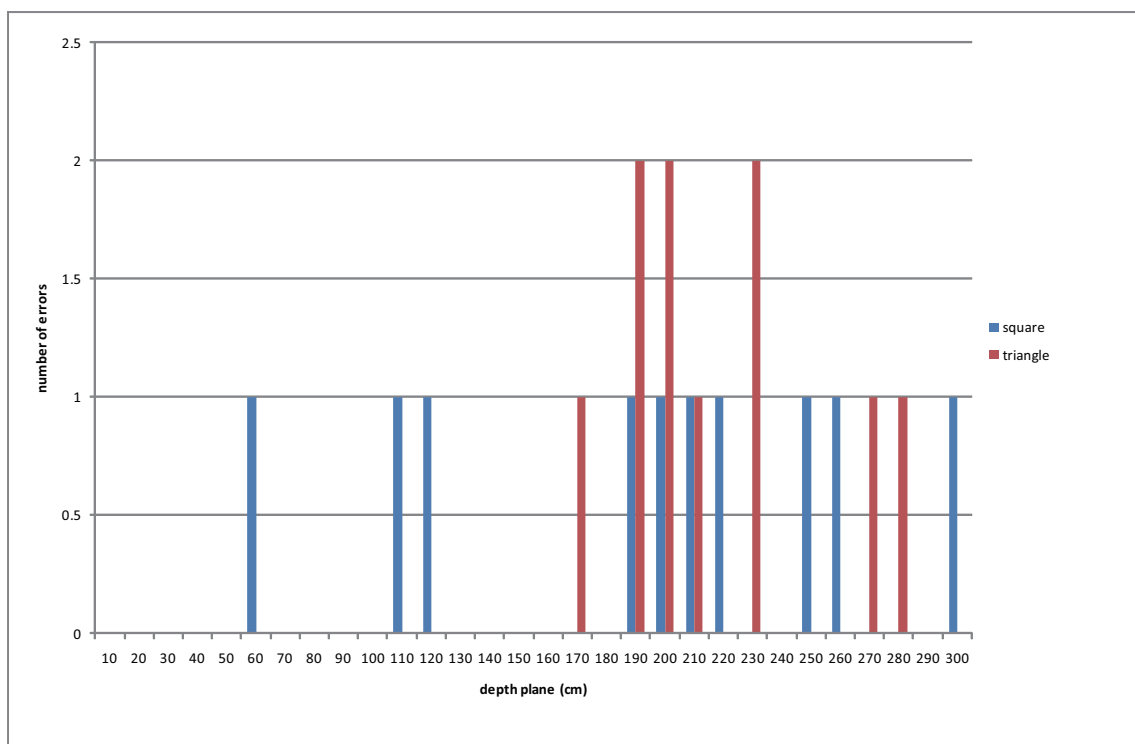


Figure 6.13: Different errors made for identifying squares and triangles on the depth planes behind the display.

6.3.5 Comparison between Square and Triangle based on time

In Front of the Display

Figure 6.14 demonstrates the different time for identifying square and triangle spent on each depth plane in front of the screen. We can observe that the differences in

the time spent between the two shapes are not large and do not have a special trend across all the depth planes.

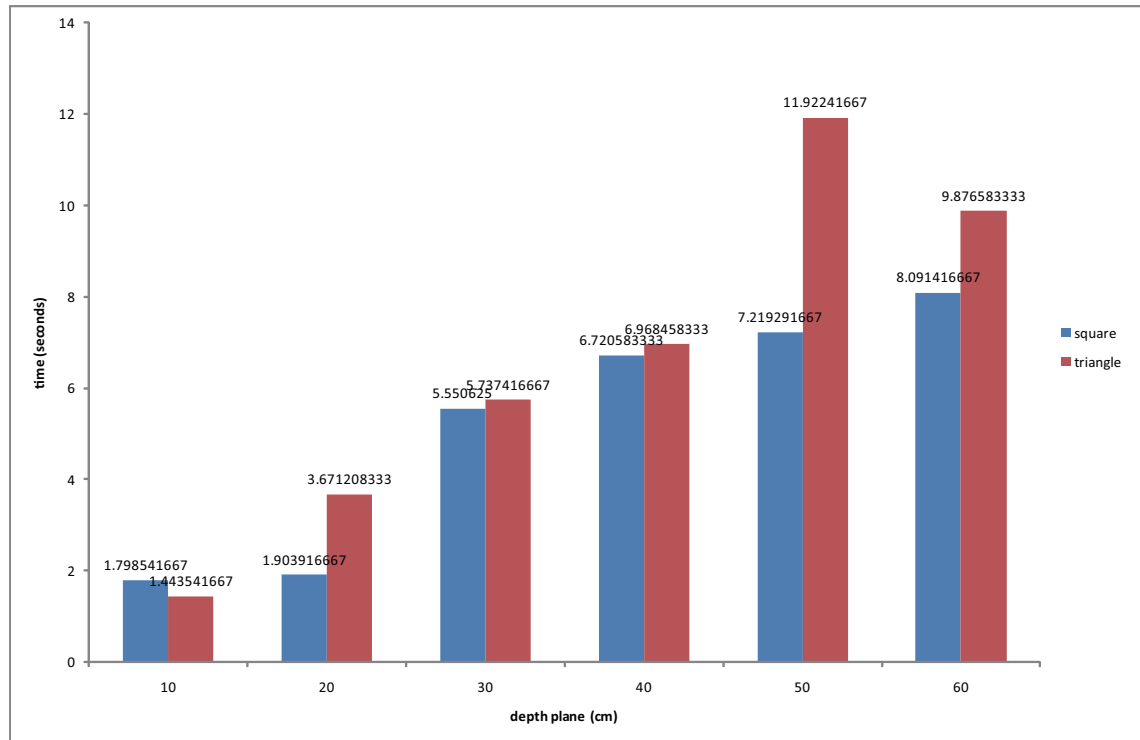


Figure 6.14: Different time spent on identifying squares and triangles on depth planes in front of the display.

Table 6.5 lists the statistical analysis results of the amount of time spent on identifying square and triangle on each depth plane. We can observe that these differences were not statistically significant. The results indicate that the shape of the object does not have an effect on the time spent. These results agree with hypothesis H5.

Behind the Display

Figure 6.15 demonstrates the different time for identifying square and triangle spent on each depth plane behind the screen. We can observe that the differences in the time spent between the two shapes are small across all the depth planes.

Table 6.6 lists the statistical analysis results of the amount of time spent on identifying square and triangle on each depth plane. No statistical significance is

Table 6.5: Time spent on square and triangle in front of the screen

Depth	Mean(s t)	StDev(s t)	95% CI	p-value	Significant?
10	(1.799 1.444)	(1.256 0.888)	(−0.201, 0.911)	0.199	No
20	(1.904 3.671)	(1.002 4.861)	(−3.770, 0.235)	0.081	No
30	(5.55 5.74)	(7.04 7.79)	(−2.28, 1.90)	0.855	No
40	(6.72 6.97)	(5.36 7.26)	(−2.46, 1.96)	0.819	No
50	(7.22 11.92)	(5.57 11.21)	(−9.68, 0.27)	0.063	No
60	(8.09 9.88)	(6.64 7.10)	(−5.13, 1.56)	0.281	No

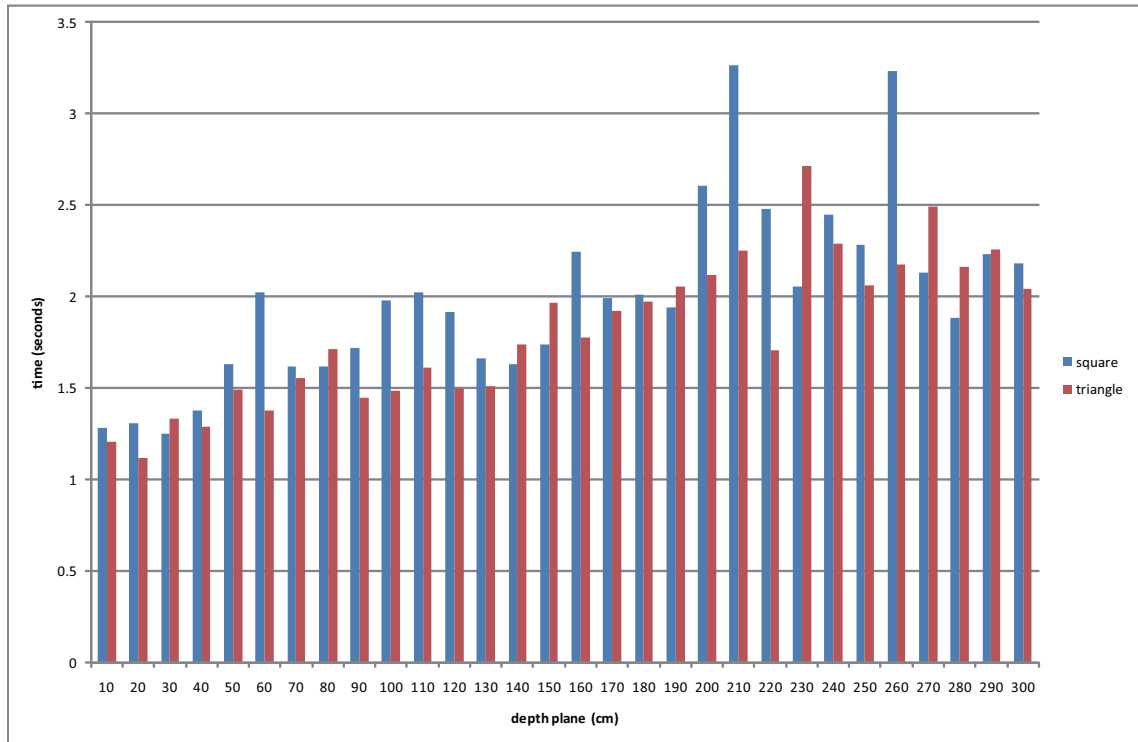


Figure 6.15: Different time spent on identifying squares and triangles on depth planes behind the display.

revealed on most of the depth plane with the exceptions on the depth planes of 20 and 220 cm. These results agree with hypothesis H6.

6.4 Conclusion

This chapter described a method that can efficiently and effectively identify the limits of the 3D Comfort Zone on a given desktop 3D display. This method was developed by using the FBO and the Stencil Buffer in OpenGL and can be easily applied to other types of 3D display systems including 3D mobile, 3DTV, and 3D cinema with trivial technical adjustments (different configurations of the cameras, display parameters, viewing distances, etc).

From the statistical analysis conducted based on both mistakes made and time spent in identifying the hidden 3D shape on depth planes in front of and behind the display screen, we concluded that the limit of the 3D Comfort Zone in front of the tested display screen should be around 30 cm. Regarding the 3D Comfort Zone limit behind the screen, although there was no statistical significance shown purely based on the analysis of the mistakes made, we run paired t-test to compare the time spent by the viewer to identify the hidden 3D shape on each depth plane with the one on the nearest depth plane tested. We concluded that 40 cm behind the screen should be the conservative perceived depth limit of the 3D Comfort Zone. In addition, we did not find any statistical significance between square and triangle in terms of the number of correct choices and the amount of time people spend on identifying them.

Detailed analysis of the meaning of our results in stereoscopic cinematography is presented in the following chapter.

Table 6.6: Time spent on square and triangle behind the screen

Depth	Mean(s t)	StDev(s t)	95% CI	p-value	Significant?
10	(1.283 1.205)	(0.734 0.659)	(−0.343, 0.499)	0.705	No
20	(1.306 1.114)	(0.473 0.492)	(−0.000, 0.384)	0.050	Yes
30	(1.250 1.335)	(0.504 0.604)	(−0.328, 0.158)	0.478	No
40	(1.379 1.288)	(0.492 0.592)	(−0.187, 0.368)	0.507	No
50	(1.627 1.490)	(0.978 1.005)	(−0.235, 0.509)	0.453	No
60	(2.020 1.380)	(2.738 0.365)	(−0.508, 1.789)	0.260	No
70	(1.618 1.553)	(1.133 1.105)	(−0.530, 0.659)	0.825	No
80	(1.617 1.715)	(0.944 2.494)	(−1.185, 0.988)	0.853	No
90	(1.716 1.448)	(1.233 0.500)	(−0.247, 0.783)	0.293	No
100	(1.977 1.482)	(1.588 0.569)	(−0.172, 1.161)	0.138	No
110	(2.024 1.613)	(2.313 1.170)	(−0.639, 1.462)	0.427	No
120	(1.917 1.495)	(1.410 0.724)	(−0.161, 1.004)	0.148	No
130	(1.659 1.509)	(0.668 0.897)	(−0.144, 0.443)	0.302	No
140	(1.631 1.735)	(0.759 0.886)	(−0.398, 0.191)	0.474	No
150	(1.740 1.966)	(0.692 1.425)	(−0.753, 0.301)	0.384	No
160	(2.245 1.776)	(1.465 0.891)	(−0.172, 1.111)	0.144	No
170	(1.994 1.920)	(1.125 1.485)	(−0.415, 0.563)	0.758	No
180	(2.011 1.971)	(1.161 1.036)	(−0.334, 0.415)	0.825	No
190	(1.943 2.057)	(0.981 1.626)	(−0.846, 0.618)	0.751	No
200	(2.605 2.118)	(3.498 2.028)	(−1.148, 2.121)	0.544	No
210	(3.267 2.248)	(3.408 1.645)	(−0.247, 2.285)	0.110	No
220	(2.478 1.708)	(1.882 0.544)	(0.040, 1.500)	0.040	Yes
230	(2.052 2.712)	(1.354 2.677)	(−1.524, 0.203)	0.127	No
240	(2.445 2.286)	(2.254 1.450)	(−0.541, 0.859)	0.642	No
250	(2.281 2.062)	(1.181 1.409)	(−0.400, 0.838)	0.472	No
260	(3.234 2.176)	(4.774 2.101)	(−1.05, 3.16)	0.309	No
270	(2.133 2.489)	(1.308 2.121)	(−1.361, 0.650)	0.472	No
280	(1.882 2.160)	(0.633 1.415)	(−0.729, 0.174)	0.216	No
290	(2.233 2.255)	(1.444 1.175)	(−0.721, 0.676)	0.948	No
300	(2.182 2.044)	(1.453 0.814)	(−0.393, 0.668)	0.597	No

Chapter 7

Discussion

This chapter compares our results with previous works, discusses our findings in terms of their meaning in S3D cinematography, describes applications and limitations of this study, and finally, establishes a basic guideline for stereoscopic cinematography. We anticipate that this guideline can help ensure a compelling and comfortable S3D viewing experience.

7.1 Limits of the 3D Comfort Zone

Table 7.1 summarises the results on the limits of the 3D Comfort Zone. Note that these studies were conducted on different 3D display systems with different viewing distances.

As shown in Table 7.1, the limits of the 3D Comfort Zone vary significantly among 3D display systems. Various factors could result in different limits of the 3D Comfort Zone, such as different 3D display systems and different experimental conditions (the size of the display, the viewing distance, etc.). We conclude that there is no consensus on one quantitative range of the 3D Comfort Zone. A recently published study conducted independently by Takashi *et al* agrees with this conclusion [Shibata et al., 2011].

Hence, it is our recommendation that stereoscopic cinematographers should test the specific limits of the 3D Comfort Zone on the target 3D display system before the actual production of stereoscopic content. This was the reason that we developed a

Human factors studies	Viewing distance (mm)	Limit in front (mm)	Limit behind (mm)
Yeh and Silverstein's	660	49	50
Williams and Parrish's	483	122	290
	965	241	579
	1448	361	869
Woods <i>et al</i> 's	800	107	145
Jones <i>et al</i> 's	700	50	60
Our results	750	300	3000

Table 7.1: Results on the limits of the 3D Comfort Zone

method that is capable of identifying the limits of the 3D Comfort Zone effectively on any given stereoscopic display system.

Another interesting point yielded from our results was that viewers were able to cope with much larger perceived depth behind the display than the depth in front of the display. Statistical results revealed that subjects could not easily perceive depth that was farther than 300mm in front of the display screen whilst no statistical significance was shown on depth perception behind the display even between 100mm the closest perceived depth plane tested and 3000mm, the farthest perceived depth plane tested. Our results agree with previous studies on limiting perceived depth in front of the display. However, it is our conclusion that large stereoscopic depth can be employed behind the display screen without causing any depth perception problem on modern desktop stereoscopic display systems (as long as there is no appearance of the eye divergence). Cinematographers should have a great deal of freedom in exploring the space behind the screen surface. The 3D research group at SONY Computer Entertainment Europe agrees with this conclusion [Benson, 2011].

Although our experiment was only conducted on static stereoscopic scenes, it is reasonable to assume that this is also the case in dynamic stereoscopic scenes as literature has shown that viewers can perceive larger depth in stereoscopic motion pictures than in static 3D images before the depth becomes excessive [Speranza

et al., 2006].

Note that an interesting point from Mendiburu is that the perceived depth behind a cinema screen is not recommended as viewers sit far away from the screen and even a small amount of depth behind the screen can cause the eye divergence [Mendiburu, 2009].

7.2 Dynamic Depth Mapping vs. Fixed Depth Mapping

Our new dynamic depth mapping approach, described in Chapter 5, was developed on top of the fixed depth mapping approach with an additional function that ensures the perceived depth on each side of the display screen remains constant despite of any change of the scene depth, eliminating the appearance of inadequate perceived depth and significantly reducing the occurrence of the sudden and dramatic perceived depth change. The experiment presented in Chapter 5 confirmed that our new dynamic depth mapping method did have an advantage over the fixed depth mapping method in controlling perceived depth in dynamic stereoscopic scenes.

Figure 7.1 demonstrates the differences between our dynamic depth mapping method and the fixed depth mapping method. Considering Figure 7.1(a), the fixed mapping method maps the near and far limits of the scene boundary onto the near and far limits of the perceived depth range specified by human factors studies. A spaceship flies from A to D. When the spaceship is at A, the farthest spot of the scene, the perceived depth reaches its maximum value behind the screen; when the spaceship flies to B, a spot close to the *ZDP*, the perceived depth dramatically decreases, $P_b \ll P_a$, as a result of substantial scene depth reduction combining with a fixed camera separation. The same effect takes place in depth perception in front of the screen as well, $P_c \ll P_d$. This is the reason that the perceived depth can become inadequate with a small scene depth using the fixed depth mapping method. Moreover, when the spaceship promptly moves from A to B, the scene depth changes abruptly and dramatically triggering a sudden and significant perceived depth change on the 3D display.

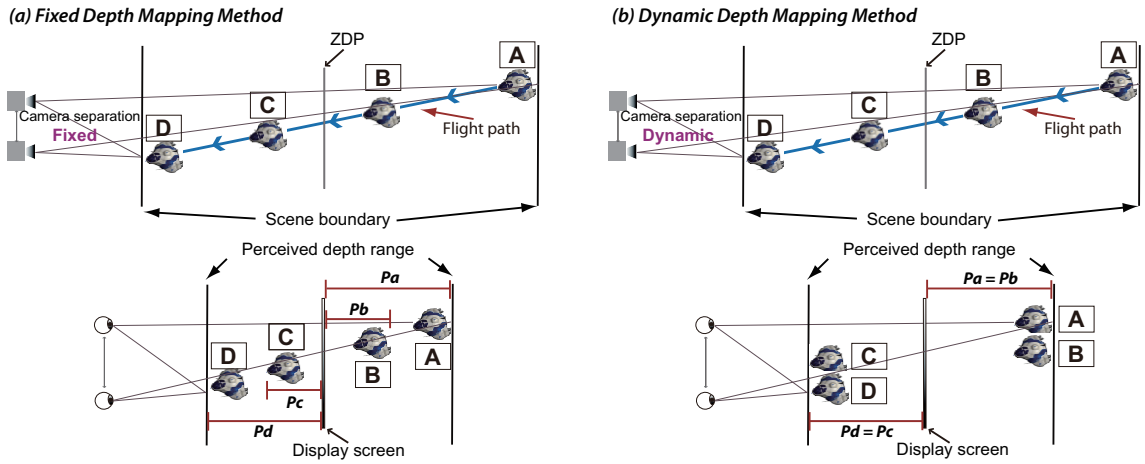


Figure 7.1: Fixed Depth Mapping vs. Dynamic Depth Mapping. In fixed depth mapping (a), the camera separation is fixed. The perceived depth changes with the position of the spaceship and can become inadequate (at B or C). In dynamic depth mapping (b), the camera separation dynamically changes with the position of the spaceship. The perceived depth on each side of the display stays constant and always occupies the whole volume of the predefined range.

With our new dynamic depth mapping method illustrated in Figure 7.1(b), the perceived depth, on each side of the display screen, is kept constant through automatic camera separation adjustments regardless of how impetuously and dramatically the scene depth changes [Sun and Holliman, 2009]. The perceived depth of the spaceship at A/D is equal to the perceived depth of the spaceship at B/C, $P_a = P_b$ and $P_d = P_c$. As the perceived depth remains unchanged (on each side of the screen) and always matches the maximum depth value of the perceived depth range, the occurrence of inadequate perceived depth is prevented, so is the sudden and dramatic change of the perceived depth on the same side of the display. The sudden and dramatic change in the sign of the scene depth and consequently the perceived depth, e.g., the spaceship rapidly flies from A to D, should simply be avoided by cinematographers [Ware et al., 1998, Speranza et al., 2006].

Note that the spaceship only moves laterally on the display plane as the spaceship moves from A/C to B/D in Figure 7.1(b). It was drawn this way only to demonstrate the constancy of the perceived depth. Viewers can still see the change of size and perspective caused by the movement of the spaceship with our new approach.

7.3 Depth of Field Blur Simulation

Figure 7.2 illustrates a dynamic scene blurred by two different DoF blur simulations. In Figure 7.2(a), the DoF is fixed and only objects coming inside the DoF are clear and seen in full sharpness. Note how the degree of blurring increases with the distance to the DoF. Figure 7.2(b) demonstrates the effect of the dynamic DoF blur simulation. A shallow DoF is created and synchronised with the flying spaceship so the spaceship is always in focus and clear. Note that an asteroid also comes into focus as the spaceship flies by it. It is a standard storytelling technique in cinematography where cinematographers place a shallow DoF upon objects on different planes, selectively redirecting audiences' attention by shifting the DoF from objects on one plane to objects on another [Sijll, 2005].

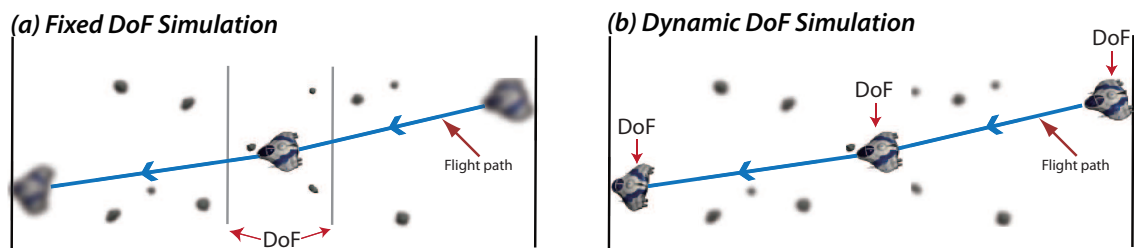


Figure 7.2: Depth of Field Blur Simulation. (a) the DoF is fixed and only objects coming inside the DoF are seen in full sharpness. The degree of blurring increases with the distance to the DoF. (b) A shallow DoF is synchronised with the spaceship so the spaceship is always in focus. An asteroid also comes into focus as the spaceship flies by it.

We investigated both Fixed and Dynamic DoF blur simulation techniques alongside with the Fixed and Dynamic Depth Mapping methods through a human-based experiment (described in Chapter 5). The results of the experiment indicated that viewers do not desire viewing blurred stereoscopic content and they much prefer depth mapping methods over DoF blur simulation techniques (fixed and dynamic).

Regarding the fixed DoF blur simulation, only one subject out of seventeen rated it above “Fair” which corresponded with the score 50 out of 100. The other sixteen all stated that they did not like this approach as they would like to see the whole context when viewing dynamic stereoscopic scenes. Two of them said that they did

not mind blurring the background objects much, it was the foreground blur that really annoyed them. This method scored the lowest mean 31 which fell into the category “Poor” in the ITU’s grading scale. The mean score of the Fixed Depth Mapping method was 68.441 and the mean score of the Dynamic Depth Mapping method was 71.206.

For the Dynamic DoF blur simulation, its mean was a little bit higher than “Fair”, 52.882. The dynamic DoF had a large standard deviation, 24.223 which was the highest of all tested methods. Scores assigned by subjects varied from 7.5, the lowest score for all methods, to 90. People disliked it for the same reason as they disliked the Fixed DOF. Those who really liked it expressed that they spontaneously focused on the flying spaceship and believed this method was a good simulation of natural vision, although they were not instructed to specifically focus on the spaceship.

Despite that applying the DoF blur simulation on stereoscopic content can diminish the Vergence-Accommodation conflict [Ukai and Howarth, 2008, Ronfard and Taubin, 2010], it is our conclusion that 3D cinematographers should not use a fixed DoF to blur the object which creates the excessive perceived depth, but rely on the depth mapping method to control the perceived depth inside a limited range. Moreover, a dynamic DoF does not work as effectively for 3D cinematographers as it works for the 2D cinematographer in selectively shifting the audience’s attention, since the viewer’s eye movement is more widely spread and viewers are less likely to follow the dominant factor in viewing stereoscopic movies than in viewing 2D movies [Hakkinen et al., 2010]. We conclude that the DoF blur simulation is not able to improve the 3D viewing experience without the help from a responsive and accurate Eye-Tracking device providing information on where the viewer is looking in real time.

7.4 Limitations and Applications of this Study

Regarding the experiment described in Chapter 5, we did not compare the results by gender, age, or time of the day the subject participated in the experiment due to

a relatively small data set (17 subjects). Since we had a relatively small subject age range (20 to 32), an uneven subject gender ratio (15 males and 2 females), and an uneven distribution of time of the day the subject participated in the experiment, the results might be biased.

The method described in Chapter 6 could only provide an approximate range of the 3D Comfort Zone. This method was designed to test a wide range of perceived depth planes with a fixed interval, d , between each plane tested. Therefore, the results could have an error margin of up to $+d$ (10 cm in our results). In addition, the results of the experiment conducted to evaluate this method could also be biased due to the same limitations discussed in the last paragraph. We also had a relatively small subject age range (23 to 34), an uneven subject gender ratio (15 males and 9 females), and an uneven distribution of time of the day the subject participated in the experiment.

Another limitation of this study was that we only tested our methods on desktop stereoscopic displays with relatively small viewing distances.

Note that the above-mentioned limitations were on the data size or the parameter values, and were not related to our methodology. Therefore, we believed that our findings could be applied to a wide range of 3D display systems, the reasons were:

- The method we proposed in identifying the limits of the 3D comfort Zone (described in Chapter 6) can be applied in other systems. The implementation of the RDS technology was novel. And the statistical analysis was solid. Users only need to make trivial adjustments (different configurations of the cameras, display parameters, viewing distances, etc) to apply our method on different stereoscopic display systems, ranging from 3D mobile to 3D cinema.
- The method we proposed in comparing different depth control methods (described in Chapter 5) can also be applied in other systems. The method analysed the data by paired t-test, which could provide trustworthy results.
- The dynamic depth mapping method we developed (described in Chapter 5) can be applied in other systems with trivial technical adjustments, and is likely to produce good results. Our method could precisely map any scene

depth range to any perceived depth range given the information about camera configurations, display parameters, and the viewing distance. And it will have an advantage over the fixed depth mapping method, since it retains the benefit of the fixed depth mapping method while holding its unique advantage: eliminating the appearance of inadequate perceived depth and significantly reducing the occurrence of the sudden and dramatic perceived depth change.

- Our findings on the DoF technique (described in Chapter 5) can be applied in other 3D systems. Our results showed that there was a large variance in viewers' ratings of the dynamic blur method and in general it did not work well. We suspect that this was because the viewer's focus does not necessarily coincide with the designed focus plane. This would remain to be a problem in all 3D display systems, since it is impossible for cinematographers to manipulate all viewers' focuses. Therefore, the DoF technique is likely to have mixed receptions by viewers.

7.5 A Guideline for Stereoscopic Cinematography

With the surge of stereoscopic cinema in recent years, extensive research has been carried out on the development of stereoscopic displays and applications. New stereoscopic technology and display systems have been invented alongside good stereoscopic imaging algorithms that can precisely control the depth mapping from the virtual space onto the physical display for static 3D images. However, relatively little research has been done in the depth perception control in stereoscopic cinematography. In this study, we performed two experiments which were designed to help 3D cinematographers better understand and control the depth perception in stereoscopic cinematography. We summarise a few well-known factors in 3D cinematography followed by our new findings (highlighted in bold text) to establish a basic guideline for stereoscopic cinematography. We anticipate that this guideline will be of particular interest not only to 3D filmmaking but also to 3D gaming, sports broadcasting, and TV production.

- Camera Configuration

The converged camera configuration, also known as the toed-in camera model, introduces undesired vertical disparity which leads to the so-called “keystone distortion” [Woods et al., 1993]. Cinematographers should try to avoid the converged camera model as much as possible and use the parallel camera arrangement to shoot/create 3D scenes [Woods et al., 1993, Diner and Fender, 1993, Ronfard and Taubin, 2010].

- Ghosting and Crosstalk

Ghosting and Crosstalk is a major contributor for visual discomfort in viewing stereoscopic contents [McAllister, 1993]. 3D cinematographers need to limit high-contrast and high-parallax stereoscopic contents to minimize those undesired effects [Lipton, 1982].

- Stereoscopic Window Violation

3D cinematographers should either move stereoscopic objects away from the borders of the screen or create a virtual floating window, the so-called “*Proscenium Arch*”, to be placed between the cameras and the 3D scene to eliminate the conflict between the stereoscopic cue of the objects and the occlusion cue of the surrounding frames of the physical viewing screen, known as the Stereoscopic Window Violation [Mendiburu, 2009, Ronfard and Taubin, 2010].

- Multi-rigging Technique

The Multi-rigging technique can enhance the 3D viewing experience by eliminating the *Cardboard* and *Puppet-Theatre* effects. Its solution is to capture the 3D scene with multiple camera rigs of different camera configurations so that an object that is far away from the cameras does not appear to be flat or lose its roundness due to the distance [Ronfard and Taubin, 2010], and at the same time an object that is close to the cameras does not appear unnaturally small.

- Limits of the 3D Comfort Zone

Our experiment agreed with the literature that there is not a consensus in the limits of the 3D Comfort Zone. Our results also indicated that viewers can

cope with much larger stereoscopic depth behind the display than the depth in front of the display. A method that can help 3D cinematographers effectively identify the limits of the 3D Comfort Zone on any given stereoscopic display was provided.

- **Depth Mapping Method**

A new depth control method was introduced. It can ensure the perceived depth on each side of the display screen remains constant despite of any change in the scene depth, eliminating the appearance of the inadequate perceived depth and the sudden and dramatic perceived depth change on the same side of the screen. Our experiment confirmed its advantage over the existing fixed depth mapping method.

- **Depth of Field Blur Simulation**

Depth of Field Blur simulation should be used by 3D cinematographers with caution. Our statistical results revealed that viewers do not find watching blurred stereoscopic content a pleasant viewing experience. We recommend cinematographers to use depth mapping methods to control the perceived depth rather than the DoF simulation.

Chapter 8

Conclusions and Future Work

8.1 Summary

This thesis aimed to improve the depth perception in stereoscopic cinematography on desktop 3D displays. 3D displays simulate the real-life viewing process by presenting different left and right images to the corresponding eyes. Comparing with traditional 2D displays, stereoscopic displays can provide viewers with a more natural viewing experience by adding a third dimension around the planar display screen. However, most of the stereoscopic display systems produce conflicting vergence and accommodation/focus cues, which require the viewer's eyes to maintain focus on the display screen yet at the same time verge (the amount of the perceived depth) away from the display screen. This so-called "Vergence-Accommodation (VA) Conflict" is regarded as a major drawback in viewing stereoscopic images. In order to reduce the visual discomfort caused by the VA Conflict, human factors studies were carried out to identify the limits of the 3D Comfort Zone on a given desktop 3D display; a new dynamic depth mapping method was developed to better control the depth perception in stereoscopic cinematography; research on investigating the practical benefits of applying Depth of Field (DOF) blur simulation on stereoscopic content was also conducted.

Human factors studies on identifying the practical limits of the 3D Comfort Zone were performed on a new method developed in C++ using OpenGL. This method was able to automatically generate and present Random Dot Stereogram

(RDS) images at different perceived depth planes both in front of and behind the 3D display screen. Statistical analysis was carried out based on mistakes made and the time spent in identifying the hidden 3D shape by viewers on each depth plane. The results were: the near limit of the 3D Comfort Zone on the given desktop 3D display should be put 30 cm in front of the screen while the far limit of the 3D Comfort Zone could be large as long as there was no appearance of the eye divergence, since no statistical significance was found even between 100mm the closest depth plane behind the screen tested and 3000mm, the farthest depth plane tested. Our results suggested different 3D systems have different limits for the 3D Comfort Zone and it is necessary to identify the individual limits of the 3D Comfort Zone on each given 3D display. The method developed is capable of defining the limits of the 3D Comfort Zone on different 3D display systems, e.g., 3D cinemas, 3D mobiles, 3DTVs, etc, with only trivial technical adjustments required. This method was considered a major contribution of this thesis.

Our new dynamic depth mapping method can adjust the camera separation according to the change of the scene depth by utilising the Frame Buffer Object (FBO) and Z-Buffer in OpenGL. The perceived depth on each side of the display screen was kept constant and always occupied the whole volume of the perceived depth range defined by human factors studies. This method was able to eliminate the dramatic and sudden change of the perceive depth while ensure a compelling stereoscopic effect without any viewing discomfort. Alongside this new approach, another four existing stereoscopic depth control methods were evaluated in a human-based trial conducted by the rules of the Recommendation ITU-R BT.500-11 [Union, 2002]. Viewers were asked to rate the quality (on a scale of 0 to 100) of the stereoscopic video sequences generated by different perceived depth control methods in OpenGL and 3D Studio Max. The tested stimuli was composed of eleven still asteroids and one moving spaceship flying back and forth through the asteroids. The dynamic depth mapping method was the most preferable choice to control the depth perception and did have statistical significance over the fixed depth mapping method developed by Jones *et al* [Jones et al., 2001]. This new dynamic depth mapping approach was considered another major contribution of this study.

The DOF blur simulation technique was also investigated in the above-mentioned experiment. Two preliminary DOF simulation methods were developed: the fixed DOF approach in which a fixed focus plane with a deep DOF was employed so that only objects residing on the focus plane were seen in full sharpness; the degree of blurring increased with the distance to the focus plane; the dynamic DOF method where a dynamic focus plane with a shallow DOF was created to follow the flying spaceship so that the spaceship was always in focus as well as objects coming within the vicinity of the spaceship. The video sequence generated by the fixed DOF method had the lowest mean score (31 out of 100) among all methods tested and most of the subjects rated it as “Poor”; the movie generated by the dynamic DOF approach had the highest standard deviation: 24.223 and only those subjects who spontaneously followed the flying spaceship liked the dynamic DOF method. The results of the DOF technique suggested that viewers are less likely to spontaneously focus on the dominant factor in viewing stereoscopic content, which was agreed by the study of Hakkinen *et al* [Hakkinen et al., 2010]. Without the help from a accurate and fast head tracker providing information about what the viewer is actually looking at in real time, the DOF blur simulation can not improve the depth quality in stereoscopic cinematography. It was our recommendation that cinematographers should apply the DOF blur on stereoscopic images with caution. We consider the results on the practical benefits of the DOF simulation a major contribution of this work.

A basic guideline for stereoscopic cinematography was introduced to summarise the new findings of this thesis, which were: (1) different 3D systems differ in the limits of the 3D Comfort Zone and cinematographers should identify the specific limits of the 3D Comfort Zone on the target display before the creation of 3D content; (2) our new dynamic depth mapping method can provide benefits in controlling depth perception in 3D cinematography over the fixed depth mapping method; (3) the DOF blur technique can not improve the stereoscopic viewing experience without the real time information about what the observer is looking at. In addition several well-known key factors in 3D cinematography, including correct camera configuration, ghosting and crosstalk, the stereoscopic window violation, and the multi-rigging

technique, were also presented in the guideline.

8.2 Future Work

8.2.1 Human Factors Studies

The need of defining practical limits of the 3D Comfort Zone has been widely recognised in the literature. Both this work and Takashi *et al*'s [Shibata et al., 2011] suggest that different 3D systems have different limits for the 3D Comfort Zone. Williams and Parrish [Williams and Parrish, 1990] and Takashi *et al* also assert that the limits of the 3D Comfort Zone expand with the increase of the viewing distance. In this thesis, a method that can efficiently identify the limits of the 3D Comfort Zone on a desktop 3D display was developed. However, we only tested it with static stereoscopic images. Future research concerning human factors studies on visual comfort of stereoscopic images should fall into the following two areas:

- Investigating visual comfort in viewing dynamic stereoscopic scenes. Speranza *et al* [Speranza et al., 2006] and Sun and Holliman [Sun and Holliman, 2009] have studied the visual comfort when viewing stereoscopic motion pictures. Both the studies have their limitations. The toed-in camera model used in Speranza *et al*'s study to construct the test stimulus is well known for creating the Vertical Parallax. In our paper, we tested only one cinematic storytelling technique: object movement. Research on visual comfort of dynamic stereoscopic scenes, which are constructed by different cinematography techniques and correct camera configurations, is needed.
- Identifying the 3D Comfort Zone on different types of stereoscopic displays. The method developed in this study is capable of identifying the limits of the 3D Comfort Zone on different 3D displays (ranging from 3D mobile to 3D cinema) with trivial adjustments to camera configurations, display parameters and the viewing distance.

8.2.2 Depth Mapping in Stereoscopic Cinematography

As discussed in the preceding chapter, our new dynamic depth mapping method does have advantages over the fixed depth mapping method in controlling depth perception in stereoscopic cinematography. Future development in the dynamic depth mapping method can include:

- Integrating our dynamic depth mapping method with Holliman’s Regions of Interest (ROI) algorithm [Holliman, 2004]. Holliman’s ROI algorithm allows users to subjectively partition the scene volume and assign the region of interest where the majority of the total perceived depth is allocated. Implementing our dynamic depth mapping method on top of this algorithm can provide more precise and adequate assignment of the perceived depth to the ROI.
- Dynamic depth control with the head-tracking mechanism. The depth perception in stereoscopic images can be further improved by combining depth mapping methods with the head-tracking technique. A fast and accurate head tracker is required to provide real-time information about where the viewer is looking. Holliman’s ROI algorithm with our dynamic depth mapping function can be employed to update the perceived depth allocation, ensuring ample perceived depth is assigned to the viewer’s interested region.

8.2.3 Depth of Field Blur Simulation

Although previous studies suggested that applying Depth of Field (DOF) on stereoscopic content can improve the S3D viewing experience, our study indicated adverse effects associated with blurring stereoscopic images. When viewing stereoscopic images, the eye movements of viewers are more widely distributed. Viewers are less likely to be influenced by cinematographers but choose to focus on objects at their own discretion. For example, objects with complex 3D structures and objects moving out of the screen often attract a great deal of the viewer’s attention [Hakkinen et al., 2010]. The results of our experiment showed that viewers regard seeing blurred 3D objects rather troublesome. The future work in applying the DOF technique on stereoscopic images includes:

- Integrating the DOF simulation with the head-tracking technique. The head tracker is responsible for providing real-time information about what the viewer is looking at so that objects in the viewer's focus are always clear and in full sharpness.
- Investigating more advanced DOF simulation techniques, e.g., the ray distribution approach, to define the optimum DOF, which can simulate the characteristics of the real eye's DOF.

Bibliography

- [Ber, 2010] (2010). Digital cinema in russia: Is 3D still a driver for the development of the cinema market? Paper presented at the 3D Media 2010, statistic from kinopoisk.ru.
- [HMD, 2011] (2011). http://commons.wikimedia.org/wiki/File%3ARealite_virtuelle.jpg (24th December, 2011).
- [Vit, 2011] (2011). <http://upload.wikimedia.org/wikipedia/commons/b/b7/VitaphoneDemo.jpg> (26th December, 2011).
- [Tel, 2011] (2011). <http://commons.wikimedia.org/wiki/File%3ATeleview.jpg> (26th December, 2011).
- [aut, 2011] (2011). <http://commons.wikimedia.org/wiki/File%3AAutostereogram.png> (25th December, 2011).
- [3dm, 2011] (2011). <http://www.3dmovielist.com/> (29th December, 2011).
- [Col, 2011] (2011). <http://www.colorcode3d.com/VSPaper.html> (24th December, 2011).
- [DaL, 2011] (2011). <http://www.da-lite.com/products/selecting.php> (26th December, 2011).
- [HDI, 2011] (2011). <http://www.hdi3d.com/products.html> (28th December, 2011).
- [CF3, 2011] (2011). <http://m.lg.com/uk/it-products/projectors/LG-CF3D.jsp> (24th December, 2011).

- [Dep, 2011a] (2011a). <http://www.depthq.com/projector.html> (23th December, 2011).
- [cav, 2011] (2011). <http://www.aec.at/center/proj/cavee> (15th December, 2011).
- [Per, 2011] (2011). http://actuality-medical.com/site/content/perspecta_display1-9.html (26th December, 2011).
- [Dep, 2011b] (2011b). <http://www.lightspacetech.com/index.html> (25th December, 2011).
- [HMZ, 2011] (2011). <http://www.sony.co.uk/product/head-mounted-display/hmz-t1> (15th December, 2011).
- [Sen, 2011] (2011). <http://sensics.com/technology/breakthrough.php> (20th December, 2011).
- [AG, 1999] AG, E., editor (1999). *3D Revelator, Quick Start Guide*. Aachen, Germany.
- [Agrawala et al., 1997] Agrawala, M., Beers, A. C., McDowall, I., Fröhlich, B., Bolas, M., and Hanrahan, P. (1997). The two-user responsive workbench: Support for collaboration through independent views of a shared space. *Computer Graphics*, 31(Annual Conference Series):327–332.
- [Akeley, 2004] Akeley, K. (2004). *Achieving near-correct focus cues using multiple image planes*. PhD thesis, Stanford University, Stanford, CA, USA.
- [Ando et al., 2005] Ando, T., Mashitani, K., Higashino, M., Kanayama, H., Murata, H., Funazou, Y., Sakamoto, N., Hazama, H., Ebara, Y., and Koyamada, K. (2005). Multiview image integration system for glassless 3D display. In *Stereoscopic Displays and Virtual Reality Systems XII. Proceedings of the SPIE*, volume 5664.
- [Atchison and Smith, 2000] Atchison, D. and Smith, G. (2000). *Optics of the Human Eye*. pp. 213-220. Butterworth-Heinemann, Oxford, UK, ISBN 0-7506-3775-7.

- [Autodesk, 2008] Autodesk (2008). Stereoscopic filmmaking whitepaper. http://images.autodesk.com/adsk/files/stereoscopic_whitepaper_final08.pdf (11th December, 2011).
- [Balogh, 2006] Balogh, T. (2006). The holovizio system. In *Stereoscopic Displays and Virtual Reality Systems XIII. Proceedings of the SPIE*, volume 6055.
- [Balogh et al., 2005] Balogh, T., Forgacs, T., Agocs, T., Balet, O., Bouvier, E., Bettio, F., Gobbetti, E., and Zanetti, G. (2005). A scalable hardware and software system for the holographic display of interactive graphics applications. In *EUROGRAPHICS 2005 Short Papers Proceedings*, Conference Held in Dublin, Ireland, August 2005.
- [Balogh et al., 2007] Balogh, T., Kovacs, P., and Barsi, A. (2007). Holovizio 3D display system. In *3DTV Conference*, pages 1–4.
- [Benson, 2011] Benson, S. (2011). The current state of 3D gaming on playstation 3. invited presentation from Sony Computer Entertainment Europe.
- [Bereby-Meyer et al., 1999] Bereby-Meyer, Y., Leiser, D., and Meyer, J. (1999). Perception of artificial stereoscopic stimuli from an incorrect viewing point. *Perception and Psychophysics*, 61(8):1555–1563.
- [Blakemore, 1970] Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Physiology*, 211:599–622.
- [Blohm et al., 1997] Blohm, W., Beldie, I. P., Schenke, K., Fazel, K., and Pastoor, S. (1997). Stereoscopic image representation with synthetic depth of field. *Journal of the Society for Information Display*, 5(3):307–313.
- [Bogaert et al., 2008] Bogaert, L., Meuret, Y., Giel, B. V., Murat, H., Smet, H. D., and Thienpont, H. (2008). Projection display for the generation of two orthogonal polarized images using liquid crystal on silicon panels and light emitting diodes. *Applied Optics*, 47:1535–1542.

- [Bogaert et al., 2009] Bogaert, L., Y. Meuret, B. V. G., Smet, H. D., and Thienpor, H. (2009). Design of a compact projection display for the visualization of 3-d images using polarization sensitive eyeglasses. *Journal of the Society for Information Display*, 17:603–609.
- [Brar et al., 2010] Brar, R. S., Surman, P., Sexton, I., Bates, R., Lee, W. K., Hopf, K., Neumann, F., Day, S. E., and Willman, E. (2010). Laser-based head-tracked 3D display research. *Journal of Display Technology*, 6:531–543.
- [Brubaker, 2009] Brubaker, B. (2009). 3D and 3D screen technology. *3D@Home Consortium Publication*.
- [Bruce et al., 2003] Bruce, V., Green, P. R., and Georgeson, M. A. (2003). *Visual Perception*. Psychology Press, New York, NY.
- [Cakmakci and Rolland, 2006] Cakmakci, O. and Rolland, J. (September 2006). Head-worn displays: A review. *Journal of Display Technology*, 2:199–216.
- [Campbell, 1957] Campbell, F. W. (1957). The depth of field of the human eye. *Journal of Modern Optics*, 4(4):157–164.
- [Chekhovskiy and Toshiyoshi, 2008] Chekhovskiy, A. and Toshiyoshi, H. (2008). The use of laser burst for volumetric displaying inside transparent liquid. *Japanese Journal of Applied Physics*, 47:6790–6793.
- [Chen et al., 2009] Chen, C. H., Huang, Y. P., Chuang, S. C., Wu, C. L., Shieh, H. P. D., Mphepo, W., Hsieh, C. T., and Hsu, S. C. (2009). Liquid crystal panel for high efficiency barrier type autostereoscopic three-dimensional displays. *Applied Optics*, 48:3446–3454.
- [Chinnock, 2009] Chinnock, C. (2009). 3D coming home in 2010. *3D@Home Consortium Publication*.
- [Choi et al., 2003] Choi, H., Min, S. W., Jung, S., Park, J.-H., and Lee, B. (2003). Multiple-viewing-zone integral imaging using a dynamic barrier array for three-dimensional displays. *Optics Express*, 11:927–932.

- [Chun et al., 2005] Chun, W., Napoli, J., Cossairt, O. S., Dorval, R. K., and Hall, D. M. (2005). Spatial 3-D Infrastructure: Display-Independent Software Framework, High-Speed Rendering Electronics, and Several New Displays. In *Stereoscopic Displays and Virtual Reality Systems XII. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, volume 5664.
- [Cossairt et al., 2007] Cossairt, O. S., Napoli, J., Hill, S. L., Dorval, R. K., and Favalora, G. E. (2007). Occlusion-capable multiview volumetric three-dimensional display. *Applied Optics*, 46:1244–1250.
- [Cruz-Neira et al., 1992] Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The cave: audio visual experience automatic virtual environment. *Communications of the ACM*, 35:64–72.
- [de Boer et al., 2007] de Boer, D. K. G., Hiddink, M. G. H., Sluijter, M., Willemsen, O. H., and de Zwart, S. T. (2007). Switchable lenticular based 2D/3D displays. In *Stereoscopic Displays and Virtual Reality Systems XIV. Proceedings of the SPIE*, volume 6490.
- [Deering, 2005] Deering, M. F. (2005). A photon accurate model of the human eye. *ACM Trans. Graph.*, 24(3):649–658.
- [Digest, 2007] Digest, S. (2007). The business case for digital 3D cinema exhibition.
- [Diner and Fender, 1993] Diner, D. B. and Fender, D. H. (1993). *Human engineering in stereoscopic viewing devices*. Plenum Press, New York, NY, USA.
- [Dodgson, 2002] Dodgson, N. A. (2002). Analysis of the viewing zone of multi-view autostereoscopic displays. In *Stereoscopic Displays and Virtual Reality Systems IX. Proceeding of the SPIE*, volume 4660, pages 254–265.
- [Dodgson, 2004] Dodgson, N. A. (2004). Variation and extrema of human interpupillary distance. In Woods, A. J., Merritt, J. O., Benton, S. A., and Bolas, M. T., editors, *Stereoscopic Displays and Virtual Reality Systems XI. Proceedings of the SPIE*, volume 5291 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 36–46.

- [Dodgson et al., 2000] Dodgson, N. A., J. R. Moore, S. R. L., Martin, G., and Canepa, P. (2000). A time sequential multi-projector autostereoscopic display. *Journal of the Society for Information Display*, 8:169–176.
- [Downing et al., 1996] Downing, E., Hesselink, L., Ralston, J., and Macfarlane, R. (1996). A three-color, solid-state, three-dimensional display. *Science*, 273:1185–1189.
- [Eichenlaub, 1998] Eichenlaub, J. B. (1998). Lightweight compact 2D/3D autostereoscopic LCD backlight for games, monitor, and notebook applications. In *Stereoscopic Displays and Virtual Reality Systems V. Proceedings of the SPIE*, volume 3295, pages 180–185.
- [Emoto et al., 2005] Emoto, M., Niida, T., and Okano, F. (2005). Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *Journal of Display Technology*, 1(2):328–340.
- [Engle, 2008] Engle, R. (2008). Beowulf 3D: A Case Study. In *Stereoscopic Displays and Applications XIX. Proceeding of the SPIE*, volume 6055.
- [Ezra et al., 1995] Ezra, D., Woodgate, G. J., Omar, B. A., Holliman, N. S., Harold, J., and Shapiro, L. S. (1995). New autostereoscopic display system. In Fisher, S. S., Merritt, J. O., and Bolas, M. T., editors, *Stereoscopic Displays and Virtual Reality Systems II. Proceeding of the SPIE*, volume 2409 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 31–40.
- [Faris, 1994] Faris, S. M. (1994). Novel 3D-stereoscopic imaging technology. In *Stereoscopic Displays and Virtual Reality Systems. Proceedings of the SPIE*, volume 2177.
- [Favalora et al., 2001] Favalora, G. E., Dorval, R. K., Hall, D. M., Giovinco, M., and Napoli, J. (2001). Volumetric three-dimensional display system with rasterization hardware. In *Stereoscopic Displays and Virtual Reality Systems VIII, Proc. SPIE*, volume 4297, pages 227–235.

- [Favalora et al., 2002] Favalora, G. E., Napoli, J., Hall, D. M., Dorval, R. K., Giovinco, M. G., Richmond, M. J., and Chun, W. S. (2002). 100 million-voxel volumetric display. In *Cockpit Displays IX: Displays for Defense Applications, Proceedings of SPIE*, volume 4712, pages 300–312.
- [Ferrari et al., 2009] Ferrari, V., G. Megali, Troia, E., Pietrabissa, A., and Mosca, F. (2009). A 3-d mixed-reality system for stereoscopic visualization of medical dataset. *IEEE Transactions on Biomedical Engineering*, 56:2627–2633.
- [Froner and Holliman, 2005] Froner, B. and Holliman, N. (2005). Implementating an improved stereoscopic camera model. In *in Eurographics Theory and Practice of Computer Graphics*.
- [Gadia et al., 2008] Gadia, D., Bonanomi, C., Rossi, M., Rizzi, A., and Marini, D. (2008). Color management and color perception issues in a virtual reality theater. In *Stereoscopic Displays and Applications XIX. Proceedings of the SPIE*, volume 6803.
- [Glassner, 1995] Glassner, A. (1995). *Principles of digital image synthesis*. Morgan Kaufmann.
- [Goldstein, 2002] Goldstein, E. B. (2002). *Sensation and Perception sixth edition*. Wadsworth.
- [Hakkinen et al., 2010] Hakkinen, J., Kawai, T., Takatalo, J., Mitsya, R., and Nyman, G. (2010). What do people look at when they watch stereoscopic movies? In *Electronic Imaging: Stereoscopic Displays & Applications XXII*, volume 7524.
- [Häkkinen et al., 2006] Häkkinen, J., Pölönen, M., Takatalo, J., and Nyman, G. (2006). Simulator sickness in virtual display gaming: a comparison of stereoscopic and non-stereoscopic situations. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, MobileHCI '06, pages 227–230, New York, NY, USA. ACM.
- [Halle, 1997] Halle, M. (1997). Autostereoscopic displays and computer graphics. *ACM Siggraph Computer Graphics*, 31:58–62.

- [Halle, 1994] Halle, M. W. (1994). Holographic stereograms as discrete imaging systems. In *Practical Holography, Proc. SPIE*, pages 73–84.
- [Hammond, 1924] Hammond, L. (1924). Stereo-scopic motion picture device. Patent no. 1,506,524.
- [Hariharan, 1996] Hariharan, P. (1996). *Optical holography : principles, techniques, and applications*. Cambridge Studies in Modern Optics. Cambridge University Press, Cambridge, ISBN 0521439655 (0-521-43965-5).
- [Harrold et al., 2004] Harrold, J., Wilkes, D., and Woodgate, G. (2004). Switchable 2D/3D display c solid phase liquid crystal microlens array. In *Proceedings of the IDW*, volume 11, pages 1495–1496.
- [Hassaine, 2010] Hassaine, D. (2010). *EFFICIENT RENDERING FOR THREE-DIMENSIONAL DISPLAYS*. PhD thesis, Durham University, Durham, UK.
- [Hassaine et al., 2010] Hassaine, D., Holliman, N. S., and Liversedge, S. P. (2010). Investigating the performance of path searching tasks in depth on multiview displays. *ACM Transactions on Applied Perception*, 8(1).
- [Haussler et al., 2008] Haussler, R., Schwerdtner, A., and Leister, N. (2008). Large holographic displays as an alternative to stereoscopic displays. In *Stereoscopic Displays and Virtual Reality Systems XIX. Proceeding of the SPIE*, volume 6803, page 68030M.
- [Heilig, 1962] Heilig, M. L. (1962). Sensorama simulator. U.S. Patent 3 050 870.
- [Helmholtz, 1962] Helmholtz, H. (1962). *Treatise on physiological optics*. Dover, translated from the third German edition of 1909-1911, originally published by Optical Society of America in 1924.
- [Hiruma and Fukuda, 1993] Hiruma, N. and Fukuda, T. (Dec. 1993). Accomodation response to binocular stereoscopic tv images and their viewing conditions. *SMPTE Journal*, pages 1137–1144.

- [Hodges and Davis, 1993] Hodges, L. F. and Davis, E. T. (1993). Geometric considerations for stereoscopic virtual environments. *Presence*, 2(1):34–43.
- [Hoffman et al., 2008] Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3).
- [Holliman et al., 2006] Holliman, N., Baugh, C., Frenk, C., Jenkins, A., Froner, B., Hassaine, D., Helly, J., Metcalfe, N., and Okamoto, T. (2006). Cosmic cookery: making a stereoscopic 3D animated movie. In *Stereoscopic Displays and Virtual Reality Systems XIII. Proceedings of the SPIE*, volume 6055, pages 34–45.
- [Holliman et al., 2007] Holliman, N., Froner, B., and Liversedge, S. (2007). An application driven comparison of depth perception on desktop 3D displays. In *Stereoscopic Displays and Virtual Reality Systems XIV. Proceedings of the SPIE*, volume 6490.
- [Holliman, 2004] Holliman, N. S. (2004). Mapping perceived depth to regions of interest in stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems XI. Proceedings of the SPIE*, volume 5291, pages 117–128.
- [Holliman, 2005] Holliman, N. S. (2005). The durham escience research institute. invited presentation to the GRIDPP13 conference, Durham.
- [Holliman, 2005] Holliman, N. S. (2005). Smoothing region boundaries in variable depth mapping for real-time stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems XII. Proceedings of the SPIE*, volume 5664, pages 281–292.
- [Holliman, 2005] Holliman, N. S. (2005). Visualizing e-science. invited presentation, Institute for Innovative Computing, Harvard.
- [Holliman, 2006] Holliman, N. S. (2006). Three-dimensional display systems. in “Handbook of Optoelectronics”, Vol II, Ed. J.P. Dakin and R.G.W. Brown, Taylor & Francis, ISBN 0 7503 0646 7, May 2006.

- [Holliman et al., 2011] Holliman, N. S., Dodgson, N. A., Favalora, G. E., and Pockett, L. (2011). Three-dimensional displays: A review and applications analysis. *IEEE Transactions on Broadcasting*, 57:362–371.
- [Howard and Rogers, 2002] Howard, I. P. and Rogers, B. J. (2002). *Seeing in Depth*. Oxford University Press, New York, USA.
- [Howlett, 1992] Howlett, E. M. (1992). High-resolution inserts in wide-angle head-mounted stereoscopic displays. In *Stereoscopic Displays and Virtual Reality Systems III. Proceedings of the SPIE.*, volume 1669.
- [Ideses and Yaroslavsky, 2004] Ideses, I. and Yaroslavsky, L. (2004). New methods to produce high quality color anaglyphs for 3-d visualization. In *Proc. Image Anal. Recog.*, volume 3212, pages 273–280.
- [Ideses and Yaroslavsky, 2005] Ideses, I. and Yaroslavsky, L. (2005). Three methods that improve the visual quality of colour anaglyphs. *Journal of Optics A: Pure and Applied Optics*, 7:755–762.
- [Johnson and Jacobsen, 2005] Johnson, R. B. and Jacobsen, G. A. (2005). Advances in lenticular lens arrays for visual display. In *Current Developments in Lens Design and Optical Engineering VI. Proceedings of the SPIE*, volume 5874.
- [Jones et al., 2007] Jones, A., McDowall, I., Yamada, H., Bolas, M., and Debevec, P. (2007). An interactive 360° light field display. In *ACM SIGGRAPH 2007 emerging technologies*, SIGGRAPH '07, New York, NY, USA. ACM.
- [Jones et al., 2001] Jones, G. R., Lee, D., Holliman, N. S., and Ezra, D. (2001). Controlling perceived depth in stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems VIII. Proceedings of the SPIE*, volume 4297, pages 42–53.
- [Jorke et al., 2009] Jorke, H., Simon, A., and Fritz, M. (May 2009). Advanced stereo projection using interference filters. *Journal of the Society for Information Display*, 17:407–410.
- [Judge, 1950] Judge, A. W. (1950). *Stereoscopic photography: its applications to science, industry, and education*. London : Chapman & Hall.

- [Julesz, 1971] Julesz, B. (1971). *Foundations of Cyclopean Perception*. The University of Chicago Press, Chicago.
- [Kikuta and Takaki, 2007] Kikuta, K. and Takaki, Y. (2007). Development of svga resolution 128-directional display. In *Stereoscopic Displays and Virtual Reality Systems XIV. Proceedings of the SPIE*, volume 6490.
- [Kimura et al., 2006] Kimura, H., Uchiyama, T., and Yoshikawa, H. (2006). Laser produced 3D display in the air. In *ACM SIGGRAPH 2006 Emerging technologies*, SIGGRAPH '06, New York, NY, USA. ACM.
- [Kooi and Toet, 2004] Kooi, F. L. and Toet, A. (2004). Visual comfort of binocular and 3D displays. *Displays*, 25(2-3):99–108.
- [Koppal et al., 2010] Koppal, S. J., Zitnick, C. L., Cohen, M., Kang, S. B., Ressler, B., and Colburn, A. (2010). A viewer-centric editor for stereoscopic cinema. *IEEE Computer Graphics and Applications*, 99(1).
- [Kruger et al., 1995] Kruger, W., Bohn, C. A., Frohlich, B., Schuth, H., Strauss, W., and Wesche, G. (1995). The responsive workbench: A virtual work environment. *Computer*, 28:42–48.
- [Lambooi et al., 2007] Lambooi, M. T. M., IJsselsteijn, W. A., and Heynderickx, I. (2007). Visual discomfort in stereoscopic displays: a review. In *Stereoscopic Displays and Virtual Reality Systems XIV. Proceedings of the SPIE*, volume 6490 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*.
- [Lambooji et al., 2009] Lambooji, M., Ijsselsteijn, W., Fortuin, M., and Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3).
- [Land, 1942] Land, E. H. (1942). Light-polarizing image full color. Patent no. 2,289,714.
- [Lane, 1982] Lane, B. (1982). Stereoscopic displays. In *Proceedings of the SPIE*, volume 0367.

- [Lang et al., 2010] Lang, M., Hornung, A., Wang, O., Poulakos, S., Smolic, A., and Gross, M. (2010). Nonlinear disparity mapping for stereoscopic 3D. *ACM Trans. Graph.*, 29(3):10.
- [Langlands, 1926] Langlands, N. (1926). Experiments on binocular vision. *Trans. Optical Soc.*, XXVII:4–82.
- [Lanman et al., 2010] Lanman, D., Hirsch, M., Kim, Y., and Raskar, R. (2010). Content-adaptive parallax barriers for automultiscopic 3D display. *ACM Transactions on Graphics*, 29.
- [Lee and Ra, 2006] Lee, Y. G. and Ra, J. B. (2006). Image distortion correction for lenticular misalignment in three-dimensional lenticular displays. *Optical Engineering*, 45:017007.
- [Leigh et al., 2007] Leigh, J., Johnson, A., Renambot, L., Sandin, D., DeFanti, T., Brown, M., Jeong, B., Jagodic, R., Krumholz, C., Svistula, D., Hur, H., Kooima, R., Peterka, T., Ge, J., and Falk, C. (2007). Emerging from the cave: Collaboration in ultra high resolution environments. In *Proceedings of First International Symposium on Universal Communication 2007*, Kyoto, Japan.
- [Lewis et al., 1971] Lewis, J. D., Verber, C. M., and McGhee, R. B. (1971). A true three-dimensional display. *IEEE Transactions on Electron Devices*, 18:724–732.
- [Lipton, 1982] Lipton, L. (1982). *Foundations of stereoscopic cinema*. Van Nostrand Reinhold. now available electronically at <http://www.3d.curtin.edu.au/cgi-bin/library/foundation.cgi> (27th December, 2011).
- [Lipton, 1990] Lipton, L. (1990). Liquid crystal shutter system for stereoscopic and other applications. Patent no. 4,967,268.
- [Lipton, 1997] Lipton, L., editor (1997). *Stereographics, Developers Handbook*. Stereographics Corporation.
- [Lipton, 2001] Lipton, L. (2001). The stereoscopic cinema: From film to digital projection. *SMPTE Motion Imaging Journal*, 110(9):586–593.

- [Lipton, 2007] Lipton, L. (2007). The last great innovation : The stereoscopic cinema. *SMPTE Motion Imaging Journal*, 116(11-12):518–523.
- [Liu and Hua, 2009] Liu, S. and Hua, H. (June 2009). Time-multiplexed dual-focal plane head-mounted display with a liquid lens. *Optics Letters*, 34:1642–1644.
- [Marcos et al., 1999] Marcos, S., Moreno, E., and Navarro, R. (1999). The depth-of-field of the human eye from objective and subjective measurements. *Vision Research*, 39(12):2039–2049.
- [Masaoka et al., 2006] Masaoka, K., Hanazato, A., Emoto, M., Yamanoue, H., Nojiri, Y., and Okano, F. (2006). Spatial distortion prediction system for stereoscopic images. *Journal of Electronic Imaging*, 15:3002–3014.
- [Mashitani et al., 2004] Mashitani, K., Hamagishi, G., Higashino, M., Ando, T., and Takemoto, S. (2004). Step barrier system multiview glassless 3D display. In *Stereoscopic Displays and Virtual Reality Systems XI. Proceedings of the SPIE*, volume 5291.
- [Matusik and Pfister, 2004] Matusik, W. and Pfister, H. (2004). 3D tv: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM SIGGRAPH*, pages 814–824.
- [McAllister, 1993] McAllister, D. F., editor (1993). *Stereo computer graphics: and other true 3D technologies*. Princeton University Press, Princeton, NJ, USA.
- [Mendiburu, 2009] Mendiburu, B. (2009). *3D Movie Making: Stereoscopic Digital Cinema from Script to Screen*. Focal Press, Elsevier.
- [Moore et al., 1996] Moore, J. R., Dodgson, N. A., Travis, A. R. L., and Lang, S. R. (1996). Time-multiplexed color autostereoscopic display. In *Stereoscopic Displays and Virtual Reality Systems VII. Proceedings of the SPIE.*, volume 2653.
- [Morishima et al., 1998] Morishima, H., Nose, H., Taniguchi, N., Inoguchi, K., and Matsumura, S. (1998). Rear-cross-lenticular 3D display without eyeglasses. In *Stereoscopic Displays and Virtual Reality Systems V. Proceedings of the SPIE*, volume 3295, pages 193–202.

- [Nishimura et al., 2007] Nishimura, H., Abe, T., Yamamoto, H., Hayasaki, Y., and Nishida, N. (2007). Development of 140-inch autostereoscopic display by use of full-color led panel. In *Light-Emitting Diodes: Research, Manufacturing, and Applications XI. Proceedings of the SPIE*, volume 6486.
- [Ogle, 1964] Ogle, K. N. (1964). *Researches in Binocular Vision*. Hafner Publishing Co. Ltd.
- [Okoshi, 1976] Okoshi, T. (1976). *Three-dimensional imaging techniques*. Academic Press, New York.
- [Papoulis and Pillai, 2002] Papoulis, A. and Pillai, S. U. (2002). *Probability, Random Variables and Stochastic Processes, 4th edition*. McGraw-Hill, ISBN:0-073-66011-6.
- [Pastoor, 1991] Pastoor, S. (1991). 3D-television: A survey of recent research results on subjective requirements. *Signal Processing: Image Communication*, 4(1):21–32.
- [Pastoor and Wopking, 1997] Pastoor, S. and Wopking, M. (1997). 3-d displays: A review of current technologies. *Displays*, 17:100–110.
- [Perlin et al., 2000] Perlin, K., Paxia, S., and Kollin, J. S. (2000). An autostereoscopic display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 319–326.
- [Perlin et al.,] Perlin, K., Poultney, C., Kollin, J. S., Kristjansson, D. T., and Paxia, S. Recent advances in the nyu autostereoscopic display. In *Stereoscopic Displays and Virtual Reality Systems VIII. Proceeding of the SPIE*, volume 4297, pages 196–203.
- [Peterka et al., 2008] Peterka, T., Kooima, R. L., Sandin, D. J., Johnson, A., Leigh, J., and DeFanti, T. A. (2008). Advances in the dynallax solid-state dynamic parallax barrier autostereoscopic visualization display system. *IEEE Transactions on Visualization and Computer Graphics*, 14:487–499.

- [Rolland et al., 1998] Rolland, J. P., Yoshida, A., Davis, L. D., and Reif, J. H. (1998). High-resolution inset head-mounted display. *Applied Optics*, 37:4183–4193.
- [Ronfard and Taubin, 2010] Ronfard, R. and Taubin, G., editors (2010). *Image and Geometry Processing for 3-D Cinematography*. Geometry and Computing. Springer-Verlag, Berlin, Heidelberg, Germany, ISBN 978-3-642-12391-7.
- [Saito et al., 2008] Saito, H., Kimura, H., Shimada, S., Naemura, T., Kayahara, J., Jarusirisawad, S., Nozick, V., Ishikawa, H., Murakami, T., Aoki, J., Asano, A., Kimura, T., Kakehata, M., Sasakia, F., Yashiro, H., Mori, M., Torizuka, K., and Ino, K. (2008). Laser-plasma scanning 3D display for putting digital contents in free space. In *Stereoscopic Displays and Applications XIX*, volume 6803, page 680309.
- [Sakamoto and Morii, 2006] Sakamoto, K. and Morii, T. (2006). Multiview 3D display using parallax barrier combined with polarizer. In *Advanced Free-Space Optical Communication Techniques/Applications II and Photonic Components/Architectures for Microwave Systems and Displays*, volume 6399.
- [Sandin et al.,] Sandin, D., Margolis, T., Dawe, G., Leigh, J., and Defanti, T. A. The varrier autostereographic display. In *Stereoscopic Displays and Virtual Reality Systems VIII. Proceeding of the SPIE*, volume 4297, pages 204–211.
- [Sanyo, 1997] Sanyo (1997). Sanyo announces development of non-glasses 15-inch xga 3D display. Sanyo Electric Corporation Limited.
- [Schiffman, 2000] Schiffman, H. R. (October 2000). *Sensation and Perception: An Integrated Approach, 5th edition*. Wiley, ISBN:0-471-24930-0.
- [Schwerdtner and Heidrich, 1998] Schwerdtner, A. and Heidrich, H. (1998). Dresden 3D display (D4D). In *Stereoscopic Displays and Virtual Reality Systems II. Proceedings of the SPIE*, volume 3295, pages 203–210.

- [Shibata et al., 2011] Shibata, T., Kim, J., Hoffman, D. M., and Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, 11:1–29.
- [Sjill, 2005] Sjill, J. V. (2005). *Cinematic Storytelling: The 100 Most Powerful Film Conventions Every Filmmaker Must Know*. Michael Wiese Productions, Cambridge, Massachusettes.
- [Sorensen et al., 2004] Sorensen, S. E. B., Hansen, P. S., and Sorensen, N. L. (2004). Method for recording and viewing stereoscopic images in color using multichrome filters. U.S. Patent 6 687 003.
- [Speranza et al., 2005] Speranza, F., Tam, W. J., Martin, T., Stelmach, L., and Ahn, C. (2005). Perceived smoothness of viewpoint transition in multi-viewpoint stereoscopic displays. In *Stereoscopic Displays and Virtual Reality Systems XII. Proceedings of the SPIE.*, volume 5664, pages 72–82.
- [Speranza et al., 2006] Speranza, F., Tam, W. J., Renaud, R., and Hur, N. (2006). Effect of disparity and motion on visual comfort of stereoscopic images. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6055, pages 94–103.
- [Spottiswoode et al., 1952] Spottiswoode, R., Spottiswoode, N. L., and Smith, C. (1952). Basic principles of the three-dimensional film. *Journal of the SMPTE*, 59:249–286.
- [Stanley et al., 2004] Stanley, M., Smith, M., Smith, A., Watson, P., Coomber, S., Cameron, C., Slinger, C., and Wood, A. (2004). 3D electronic holography display system using a 100-megapixel spatial light modulator. In *Optical Design and Engineering. Proceedings of the SPIE*, volume 5249.
- [Sugihara et al., 1999] Sugihara, T., Miyasato, T., and Nakatsu, R. (1999). An evaluation of visual fatigue in 3-d displays: focusing on the mismatching of convergence and accommodation. *IEICE TRANS. ELECTRON.*, E82-C(10):1814–1822.

- [Sullivan, 2004] Sullivan, A. (2004). Depthcube solid-state 3D volumetric display. In *Stereoscopic Displays and Virtual Reality Systems XI, Proc. SPIE*, volume 5291, pages 279–284.
- [Sun and Holliman, 2009] Sun, G. and Holliman, N. S. (2009). Evaluating methods for controlling depth perception in stereoscopic cinematography. In *Stereoscopic Displays and Virtual Reality Systems XX. Proceedings of SPIE-IS&T Electronic Imaging, SPIE*, volume 7237.
- [Surman et al., 2008] Surman, P., Sexton, I., Hopf, K., Lee, W. K., Buckley, E., Jones, G., and Bates, R. (2008). European research into head tracked autostereoscopic displays. In *3DTV Conference*, pages 161–164.
- [Takahashi and Hirooka, 2008] Takahashi, H. and Hirooka, S. (2008). Stereoscopic see-through retinal projection head-mounted display. In *Stereoscopic Displays and Virtual Reality Systems XIX. Proceedings of the SPIE.*, volume 6803.
- [Takaki, 2001] Takaki, Y. (2001). T. honda and y. kajiki and k. susami and t. hamaguchi and t. endo and t. hatada and t. fujii. In *SPIE Critical Reviews*, volume CR76, pages 218–249.
- [Takaki, 2005] Takaki, Y. (2005). Thin-type natural three-dimensional display with 72 directional images. In *Stereoscopic Displays and Virtual Reality Systems XII. Proceedings of the SPIE*, volume 5664.
- [Takaki, 2006] Takaki, Y. (2006). High-density directional display for generating natural three-dimensional images. *Proceedings of the IEEE*, 94:654–663.
- [Takaki, 2009a] Takaki, Y. (2009a). Super multi-view display and holographic display. In *LEOS Annual Meeting Conference Proceedings, LEOS '09. IEEE*, pages 12–13.
- [Takaki, 2009b] Takaki, Y. (2009b). Super multi-view display with 128 viewpoints and viewpoint formation. In *Stereoscopic Displays and Virtual Reality Systems XX. Proceedings of the SPIE*, volume 7237.

- [Takaki and Hayashi, 2008] Takaki, Y. and Hayashi, Y. (2008). Increased horizontal viewing zone angle of a hologram by resolution redistribution of a spatial light modulator. *Applied Optics*, 47:D6–D11.
- [Takaki and Okada, 2009] Takaki, Y. and Okada, N. (2009). Hologram generation by horizontal scanning of a high-speed spatial light modulator. *Applied Optics*, 48:3255–3260.
- [Tang and Evans, 2007] Tang, S. T. W. and Evans, B. J. W. (January 2007). The near mallett unit foveal suppression test: a cross-sectional study to establish test norms and relationship with other optometric tests. *Ophthalmic and Physiological Optics*, 27:31–43.
- [Thompson, 2005] Thompson, J. I. (2005). *A Three Dimensional Helmet Mounted Primary Flight Reference for Paratroopers*. PhD thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, USA.
- [Trayner and Orr, 1997] Trayner, D. and Orr, E. (1997). Developments in autostereoscopic displays using holographic optical elements. In *Stereoscopic Displays and Virtual Reality Systems IV. Proceedings of the SPIE*, volume 3012.
- [Trayner and Orr, 1996] Trayner, D. J. and Orr, E. (1996). Autostereoscopic display using holographic optical elements. In *Stereoscopic Displays and Virtual Reality Systems III. Proceedings of the SPIE*, volume 2653, pages 65–74.
- [Tsai et al., 2000] Tsai, C. H., Lai, P., Lee, K., and Lee, C. K. (2000). Fabrication of a large f-number lenticular plate and its use as a small-angle flat-top diffuser in autostereoscopic display screens. In *Stereoscopic Displays and Virtual Reality Systems VII. Proceedings of the SPIE*, volume 3957, pages 322–329.
- [Ukai and Howarth, 2008] Ukai, K. and Howarth, P. A. (2008). Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. *Displays*, 29(2):106–116.

- [Union, 2002] Union, I. T. (2002). Methodology for the subjective assessment of the quality of television pictures. <http://www.itu.int/rec/R-REC-BT.500-11-200206-I> (26th December, 2011).
- [Urey et al., 2011] Urey, H., Chellappan, K. V., Erden, E., and Surman, P. (April, 2011). State of the art in stereoscopic and autostereoscopic displays. In *Proceedings of the IEEE*, volume 99, pages 540–550.
- [Valyus, 1966] Valyus, N. A. (1966). *Stereoscopy*. The Focal Press.
- [van Berkel, 1999] van Berkel, C. (1999). Image preparation for 3D lcd. In *Stereoscopic Displays and Virtual Reality Systems VI. Proceedings of the SPIE.*, volume 3639.
- [van Berkel and Clarke, 1996] van Berkel, C. and Clarke, J. (1996). Characterisation and optimisation of 3D-lcd module design. In *Stereoscopic Displays and Virtual Reality Systems IV. Proceedings of the SPIE.*, volume 3012.
- [Vishwanath et al., 2005] Vishwanath, D., Girshick, A., and Banks, M. (2005). Perception of artificial stereoscopic stimuli from an incorrect viewing point. *Nature Neuroscience*, 8(10):1401–1410.
- [Walworth, 1984] Walworth, V. (1984). Three-dimensional projection with circular polarizers. *Proceeding of the SPIE. Optics in Entertainment II*, 462.
- [Wann and Mon-Williams, 1997] Wann, J. P. and Mon-Williams, M. (1997). Health issues with virtual reality displays: what we do know and what we don't. *SIGGRAPH Comput. Graph.*, 31(2):53–57.
- [Wann and Mon-Williams, 2002] Wann, J. P. and Mon-Williams, M. (2002). *Measurement of Visual Aftereffects Following Virtual Environment Exposure*. Handbook of Virtual Environments. Lawrence Erlbaum Associates , London, ISBN 080583270X.
- [Ware and Franck, 1996] Ware, C. and Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics*, 15(2):121–140.

- [Ware et al., 1995] Ware, C., Gobrecht, C., and Paton, M. (1995). Algorithm for dynamic disparity adjustment. In *Stereoscopic Displays and Virtual Reality Systems II. Proceedings of the SPIE*, volume 2409, pages 150–156.
- [Ware et al., 1998] Ware, C., Gobrecht, C., and Paton, M. A. (1998). Dynamic adjustment of stereo display parameters. *IEEE Transactions on Systems, Man and Cybernetics*, 28(1):56–65.
- [Wartell, 2001] Wartell, Z. (2001). *Stereoscopic Head-Tracker Displays: Analysis and Development of Display Algorithms*. PhD thesis, Georgia Institute of Technology.
- [Wartell et al., 1999] Wartell, Z., Hodges, L. F., and Ribarsky, W. (1999). Balancing fusion, image depth and distortion in stereoscopic head-tracked displays. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 351–358, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.
- [Watt et al., 2005a] Watt, S. J., Akeley, K., Ernst, M. O., and Banks, M. S. (2005a). Focus cues affect perceived depth. *Journal of Vision*, 5(10):834–862.
- [Watt et al., 2005b] Watt, S. J., Akeley, K., Girshick, A. R., and Banks, M. S. (2005b). Achieving near-correct focus cues in a 3D display using multiple image planes. In *Human Vision and Electronic Imaging X*, volume 5666, pages 393–401.
- [Wheatstone, 1838] Wheatstone, C. (1838). Contributions to the Physiology of Vision. Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. *Philosophical Transactions Series I*, 128:371–394.
- [Williams and Parrish, 1990] Williams, S. P. and Parrish, R. V. (1990). New computational control techniques and increased understanding for stereo 3D displays. In *Proceedings of SPIE Stereoscopic Displays and Applications*, pages 73–82, Santa Clara.
- [Woodgate et al., 1997] Woodgate, G. J., Ezra, D., Harrold, J., Holliman, N. S., Jones, G. R., and Moseley, R. R. (1997). Observer-tracking autostereoscopic

- 3D display systems. In *Stereoscopic Displays and Virtual Reality Systems IV. Proceedings of the SPIE*, volume 3012, pages 187–198.
- [Woodgate et al., 1998] Woodgate, G. J., Ezra, D., Harrold, J., Holliman, N. S., Jones, G. R., and Moseley, R. R. (1998). Autostereoscopic 3D display systems with observer tracking. *Signal Processing: Image Communication*, 14:131–145.
- [Woodgate et al., 2000] Woodgate, G. J., Harrold, J., Jacobs, A. M. S., Moseley, R. R., and Ezra, D. (2000). Flat panel autostereoscopic displays characterisation and enhancement. In *Proceedings of the SPIE*, volume 3957.
- [Woods et al., 1993] Woods, A. J., Docherty, T., and Koch, R. (1993). Image distortions in stereoscopic video systems. In *Stereoscopic Displays and Applications IV. Proceedings of the SPIE*, volume 1915, pages 36–47.
- [Woods and Rourke, 2004] Woods, A. J. and Rourke, T. (2004). Ghosting in anaglyphic stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems XI. Proceedings of the SPIE*, volume 5291.
- [Woods and Tan, 2002] Woods, A. J. and Tan, S. S. L. (2002). Characterising sources of ghosting in time-sequential stereoscopic video displays. In *Stereoscopic Displays and Virtual Reality Systems IX, Proc. SPIE*, volume 4660.
- [Wopking, 1995] Wopking, M. (1995). Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus. *Journal of the SID*, 3(3):101–103.
- [Yamamoto et al., 2002] Yamamoto, H., Kouno, M., Muguruma, S., Hayasaki, Y., Nagai, Y., Shimizu, Y., and Nishida, N. (2002). Enlargement of viewing area of stereoscopic full-color led display by use of a parallax barrier. *Applied Optics*, 41:6907–6919.
- [Yamanoue et al., 2006] Yamanoue, H., Okui, M., and Okano, F. (2006). Geometrical analysis of puppet-theater and cardboard effects in stereoscopic hdtv images. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(6):744–752.

- [Yamanoue et al., 2000] Yamanoue, H., Okui, M., and Yuyama, I. (2000). A study on the relationship between shooting conditions and cardboard effect of stereoscopic images. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(3):411–416.
- [Yano et al., 2004] Yano, S., Emoto, M., and Mitsuhashi, T. (2004). Two factors in visual fatigue caused by stereoscopic hdtv images. *Displays*, 25(4):141–150.
- [Yano et al., 2002] Yano, S., Ide, S., Mitsuhashi, T., and Thwaites, H. (2002). A study of visual fatigue and visual comfortnext term for 3D hdtv/hdtv images. *Displays*, 23(4):191–201.
- [Yeh, 1993] Yeh, Y. Y. (1993). Visual and perceptual issues in stereoscopic color displays. pages 50–70.
- [Yeh and Silverstein, 1990] Yeh, Y. Y. and Silverstein, L. D. (1990). Limits of fusion and depth judgment in stereoscopic color displays. *Hum. Factors*, 32(1):45–60.

Appendix A

Dynamic Depth Mapping Method

A.1 Initiating the Frame Buffer Object

```
GLuint fbo;                                // our handle to the FBO
GLuint depthBuffer;                        // our handle to the depth buffer

void InitFramebuffer(GLuint fbo, GLuint depthBuffer)
{
    // setup our FBO
    glGenFramebuffersEXT(1, &fbo);
    glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, fbo);

    // create the render buffer for depth
    glGenRenderbuffersEXT(1, &depthBuffer);

    // bind the depthBuffer to current render buffer
    glBindRenderbufferEXT(GL_RENDERBUFFER_EXT, depthBuffer);

    // request storage space of the size of the depth buffer
    glRenderbufferStorageEXT(GL_RENDERBUFFER_EXT, GL_DEPTH_COMPONENT,
                             bufferWidth, bufferHeight);
```

```
// attach the depth render buffer to the FBO as its depth attachment
glFramebufferRenderbufferEXT(GL_FRAMEBUFFER_EXT, GL_DEPTH_ATTACHMENT_EXT,
                             GL_RENDERBUFFER_EXT, depthBuffer);

// check the status of the FBO
GLenum status = glCheckFramebufferStatusEXT(GL_FRAMEBUFFER_EXT);
if(status != GL_FRAMEBUFFER_COMPLETE_EXT)
    exit(1);

// unbind the FBO for now
glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, 0);
}
```

A.2 Updating Scene Boundaries with the Z-Buffer

```
double sceneBoundaries_near;    // near limit of the actual scene depth
double sceneBoundaries_far;     // far limit of the actual scene depth

void UpdateSceneBoundaries(int pixelHeight,int pixelWidth,int farZ,int nearZ)
{
    // resolution in pixels
    int height = pixelHeight;
    int width  = pixelWidth;

    // distances to far and near clipping planes
    int fz = farZ, nz = nearZ;

    int numPixels = height * width;
    GLfloat *depths = new GLfloat[numPixels];

    // read the Z-Buffer value
```

```
glReadPixels(0,0,width,height,GL_DEPTH_COMPONENT,GL_FLOAT,depths);

double minPixelDepth = 10000000.0;
double maxPixelDepth = 0.1;

for(int i = 0; i < numPixels; i++)
{
    // converting the Z-buffer value to scene depth value
    GLdouble z = (depths[i] - 0.5) * 2.0;
    GLdouble pixelDepth = -2*fz*nz/(z*(fz-nz)-(fz+z));

    // replacing max and min with current scene depth values
    if (pixelDepth < minPixelDepth)
        minPixelDepth = pixelDepth;
    if ((pixelDepth > maxPixelDepth) && pixelDepth < farZ)
        maxPixelDepth = pixelDepth;
}

delete [] depths;

// update scene boundaries
sceneBoundaries_near = minPixelDepth;
sceneBoundaries_far  = maxPixelDepth;
}
```

A.3 Off-screen Rendering

```
void DrawOffScreen(GLuint fbo)
{
    // bind the FBO so we can render to it
    glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, fbo);
```

```
// save the view port and set it to the size of the buffer
glPushAttrib(GL_VIEWPORT_BIT);
glViewport(0,0,bufferWidth,bufferHeight);

DrawMonoImage();

// update the scene boudaries
UpdateSceneBoundaries();

// restore old view port and set rendering back to default frame buffer
glPopAttrib();
glBindFramebufferEXT(GL_FRAMEBUFFER_EXT, 0);
}
```

Appendix B

Generating RDS Images with the Stencil Buffer

B.1 Drawing a RDS image

```
GLubyte RDSImage[imageHeight][imageWidth][3];

void DrawRDSImage(int imageHeight,int imageWidth,GLubyte RDSImage,int dotSize)
{
    int dotColour;          // variable to store the colour of the dot
    int i, j, k, n, m, x;   // variables for the loops

    for (i=0; i<(imageHeight/dotSize); i++)
    {
        k = dotSize * i;
        for (j=0; j<(imageWidth/dotSize); j++)
        {
            n = j * dotSize;

            // ranNum is between 1 and 100
            int ranNum = rand() % 100 + 1;
            if (ranNum<=50)
```

```

        dotColour = 255; // dot colour is white
    else
        dotColour = 0;   // dot colour is black

    for (m=0; m<dotSize; m++)
    {
        for (x=0; x<dotSize; x++)
        {
            // assign colour to each dot
            RDSImage[k+m][n+x][0] = (GLubyte) dotColour;
            RDSImage[k+m][n+x][1] = (GLubyte) dotColour;
            RDSImage[k+m][n+x][2] = (GLubyte) dotColour;
        }
    }
}
}
}

```

B.2 Rendering Stereoscopic RDS images with the Stencil Buffer

```

void RenderStereoRDSImage(int frameNO)
{
    if (frameNO == 0)                // render the left frame
        DrawRDSImage();
    else
    {
        DrawRDSImage();
        glTranslated(1.0,0.0,0.0); // horizontally shift the viewing matrix

        // enable Stencil Buffer and set it to 1 where we draw the pattern
    }
}

```

```
    glEnable(GL_STENCIL_TEST);
    glStencilFunc(GL_ALWAYS,1,1);
    glStencilOp(GL_KEEP,GL_KEEP,GL_REPLACE);

    DrawPattern();

    // draw only where Stencil Buffer is 1 and keep it unchanged
    glStencilFunc(GL_EQUAL,1,1);
    glStencilOp(GL_KEEP,GL_KEEP,GL_KEEP);

    // redraw the RDS image
    DrawRDSImage();
}
}
```